

p. 325

The Telecommunications and Data Acquisition Progress Report 42-108

October-December 1991

E. C. Posner
Editor

(NASA-CR-190155) THE TELECOMMUNICATIONS AND
DATA ACQUISITION REPORT Progress Report,
Oct. - Dec. 1991 (JPL) 325 p CSCL 17B

N92-24301
--THRU--
N92-24323
Unclas
0079860

G3/32

February 15, 1992



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California



The Telecommunications and Data Acquisition Progress Report 42-108

October–December 1991

E. C. Posner
Editor

**ORIGINAL CONTAINS
COLOR ILLUSTRATIONS**

February 15, 1992



National Aeronautics and
Space Administration

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

The research described in this publication was carried out by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.

Preface

This quarterly publication provides archival reports on developments in programs managed by JPL's Office of Telecommunications and Data Acquisition (TDA). In space communications, radio navigation, radio science, and ground-based radio and radar astronomy, it reports on activities of the Deep Space Network (DSN) in planning, in supporting research and technology, in implementation, and in operations. Also included is standards activity at JPL for space data and information systems and reimbursable DSN work performed for other space agencies through NASA. The preceding work is all performed for NASA's Office of Space Operations (OSO). The TDA Office also performs work funded by two other NASA program offices through and with the cooperation of the Office of Space Operations. These are the Orbital Debris Radar Program (with the Office of Space Station) and 21st Century Communication Studies (with the Office of Aeronautics and Exploration Technology).

In the search for extraterrestrial intelligence (SETI), *The TDA Progress Report* reports on implementation and operations for searching the microwave spectrum. In solar system radar, it reports on the uses of the Goldstone Solar System Radar for scientific exploration of the planets, their rings and satellites, asteroids, and comets. In radio astronomy, the areas of support include spectroscopy, very long baseline interferometry, and astrometry. These three programs are performed for NASA's Office of Space Science and Applications (OSSA), with the Office of Space Operations funding DSN operational support.

Finally, tasks funded under the JPL Director's Discretionary Fund and the Caltech President's Fund that involve the TDA Office are included.

This and each succeeding issue of *The TDA Progress Report* will present material in some, but not necessarily all, of the following categories:

OSO Tasks:

- DSN Advanced Systems
 - Tracking and Ground-Based Navigation
 - Communications, Spacecraft-Ground
 - Station Control and System Technology
 - Network Data Processing and Productivity
- DSN Systems Implementation
 - Capabilities for Existing Projects
 - Capabilities for New Projects
 - New Initiatives
 - Network Upgrade and Sustaining
- DSN Operations
 - Network Operations and Operations Support
 - Mission Interface and Support
 - TDA Program Management and Analysis
- Ground Communications Implementation and Operations
- Data and Information Systems
- Flight-Ground Advanced Engineering
- Long-Range Program Planning

OSO Cooperative Tasks:

- Orbital Debris Radar Program
- 21st Century Communication Studies

OSSA Tasks:

Search for Extraterrestrial Intelligence

Goldstone Solar System Radar

Radio Astronomy

Discretionary Funded Tasks

Contents

OSO TASKS DSN Advanced Systems TRACKING AND GROUND-BASED NAVIGATION

A Demonstration of Centimeter-Level Monitoring of Polar Motion With the Global Positioning System	15
U. J. Lindqwister, A. P. Freedman, and G. Blewitt	
NASA Code 310-10-61-87-02	
A High-Performance Hg⁺ Trapped Ion Frequency Standard	10
J. D. Prestage, R. L. Tjoelker, G. J. Dick, and L. Maleki	
NASA Code 310-10-62-15-00	
Criteria for Ultrastable Operation of the Trapped Ion Frequency Standard	19
R. L. Tjoelker, J. D. Prestage, and L. Maleki	
NASA Code 310-10-62-15-00	
Orbit-Determination Performance of Doppler Data for Interplanetary Cruise Trajectories	
Part I: Error Analysis Methodology	31
J. S. Ulvestad and S. W. Thurman	
NASA Code 310-10-63-50-00	
Orbit-Determination Performance of Doppler Data for Interplanetary Cruise Trajectories	
Part II: 8.4-GHz Performance and Data-Weighting Strategies	49
J. S. Ulvestad	
NASA Code 310-10-63-50-00	
The X-Windows Interactive Navigation Data Editor	66
G. C. Rinker	
NASA Code 310-10-63-53-00	

COMMUNICATIONS, SPACECRAFT-GROUND

Alignment of the Atmospheric Visibility Monitoring Telescope	79
K. Cowles and D. Erickson	
NASA Code 310-20-67-88-03	
Nonlinear, Nonbinary Cyclic Group Codes	84
G. Solomon	
NASA Code 310-30-71-83-02	

STATION CONTROL AND SYSTEM TECHNOLOGY

Real-Time Antenna Fault Diagnosis Experiments at DSS 13	96
J. Mellstrom, C. Pierson, and P. Smyth	
NASA Code 310-20-65-91-00	
Pointing-Error Simulations of the DSS-13 Antenna Due to Wind Disturbances	109
W. Gawronski, B. Bienkiewicz, and R. E. Hill	
NASA Code 310-20-65-63-00	
DSS-13 Beam-Waveguide Antenna Performance in the Bypass Mode	135
S. R. Stewart	
NASA Code 310-30-69-91-12	

The Electrical Conductivities of the DSS-13 Beam-Waveguide Antenna Shroud Material and Other Antenna Reflector Surface Materials	154
T. Y. Otoshi, M. M. Franco, and H. F. Reilly, Jr.	
NASA Code 310-20-65-86-08	

NETWORK DATA PROCESSING AND PRODUCTIVITY

Software-Engineering Process Simulation Model (SEPS)	165
C. Y. Lin, T. Abdel-Hamid, and J. S. Sherif	
NASA Code 310-40-73-91-62	

DSN Systems Implementation CAPABILITIES FOR EXISTING PROJECTS

Node Synchronization Schemes for the Big Viterbi Decoder	186
K.-M. Cheung, L. Swanson, and S. Arnold	
NASA Code 314-30-51-01-17	

250-kW CW Klystron Amplifier for Planetary Radar	201
R. Cormier and A. Mizuhara	
NASA Code 314-30-63-31-03	

An Assembly to Provide a Time Reference for the JPL Network Operations Control Center, Real-Time	214
P. A. Clements, J. G. Lopez, and M. J. Kelley	
NASA Code 314-30-50-40-18	

NETWORK UPGRADE AND SUSTAINING

Dynamic Modeling of the Servovalves Incorporated in the Servo Hydraulic System of the 70-Meter DSN Antennas	222
R. D. Bartos	
NASA Code 314-30-42-10-87	

A Microwave Holography Methodology for Diagnostics and Performance Improvement for Large Reflector Antennas	235
D. J. Rochblatt	
NASA Code 314-30-43-10-31	

Performance Improvement of DSS-13 34-Meter Beam-Waveguide Antenna Using the JPL Microwave Holography Methodology	253
D. J. Rochblatt and B. L. Seidel	
NASA Code 314-30-43-10-31	

DSN Operations TDA PROGRAM MANAGEMENT AND ANALYSIS

DSN <i>G/Top</i> and Telecommunications System Performance	271
C. Stelzried, R. Clauss, W. Rafferty, and S. Petty	
NASA Code 314-40-31-30-06	

OSC Cooperative Tasks TWENTY-FIRST CENTURY COMMUNICATION STUDIES

Position Determination of a Lander and Rover at Mars With Earth-Based Differential Tracking	279
R. D. Kahn, W. M. Folkner, C. D. Edwards, and A. Vijayaraghavan	
NASA Code 316-30-19-41-07	

OSSA Tasks
SEARCH FOR EXTRATERRESTRIAL INTELLIGENCE

Using Single Buffers and Data Reorganization to Implement a Multi-Megasample Fast Fourier Transform	294
R. D. Brown NASA Code 108-30-30-40-01	
Author Index, 1991	307

51-04
778-1
P-9
N92-24302
JJ574450

A Demonstration of Centimeter-Level Monitoring of Polar Motion With the Global Positioning System

U. J. Lindqwister, A. P. Freedman, and G. Blewitt
Tracking Systems and Applications Section

Daily estimates of the Earth's pole position have been obtained with the Global Positioning System (GPS) by using measurements obtained during the GIG'91 experiment from January 22 to February 13, 1991. Data from a globally distributed network consisting of 21 Rogue GPS receivers were chosen for the analysis. A comparison of the GPS polar motion series with nine 24-hour very long baseline interferometry (VLBI) estimates yielded agreement in the day-to-day pole position of about 1.5 cm for both X and Y polar motion. A similar comparison of GPS and satellite laser ranging (SLR) data showed agreement to about 1.0 cm. These preliminary results indicate that polar motion can be determined by GPS independent of, and at a level comparable to, that which can be obtained from either VLBI or SLR. Furthermore, GPS can provide these data with a daily frequency that neither alternative technique can readily achieve. Thus, GPS promises to be a powerful tool for determining high-frequency platform parameter variations, essential for the ultraprecise spacecraft-tracking requirements of coming years.

I. Introduction

Estimating variations in polar motion, the direction of the Earth's rotation vector with respect to a terrestrial reference frame, has until recently been the province of two space geodetic techniques: very long baseline interferometry (VLBI) and satellite laser ranging (SLR). Reported accuracies from these techniques are typically at the level of 0.3–0.7 mas (1–2 cm),¹ and these techniques generate either 24-hr estimates every few days (VLBI) or 3-day averaged estimates every 3 days (SLR) [1,2]. Requirements by the DSN for pole-position accuracy (one to two weeks after the fact) range from about 5 cm (1 σ error)

for the TOPEX mission to under 3 cm for such missions as CRAF/Cassini. Although the non-DSN VLBI data sets that are currently utilized to provide non-real-time polar motion can nominally achieve these goals, two problems may arise. First, the dependence on non-DSN data collection and processing exposes the DSN to the problems, either technical or financial, which these techniques may encounter. Second, the 3–5 day frequency of these other techniques is inadequate to catch rapid variations in polar motion, which still have not been adequately studied.

The Global Positioning System (GPS) has been proposed as an alternative high-precision technique for monitoring pole position at more frequent intervals than those of VLBI or SLR [3,4]. An earlier study [5] demonstrated

¹ Polar motion values are usually given in milliarcseconds (mas), where 1 mas \approx 5 nrad \approx 3 cm at the Earth's surface.

the potential of using GPS to measure polar motion with a sparse global station network and a limited satellite constellation consisting of six GPS satellites. The GIG'91 GPS experiment, performed in early 1991 under the auspices of the International Earth Rotation Service (IERS), employed a large number of GPS receivers distributed worldwide and utilized the 15 available GPS satellites. That experiment provided the first opportunity to test GPS as a competitive technique for monitoring Earth rotation and polar motion [6]. This article presents GPS estimates of the Earth's pole position and its day-to-day variability, and compares these results with those from other space geodetic techniques. The high quality and daily frequency demonstrated by these measurements argue that GPS should be implemented as a standard technique for monitoring high-frequency polar motion variations in support of spacecraft navigation by the DSN. GPS would then serve as a complement to current VLBI techniques, which would remain to provide necessary long-term reference-frame stability and redundancy.

II. Analysis

The first GPS IERS and Geodynamics (GIG'91) experiment was carried out over a three-week period from January 22 to February 13, 1991. It involved numerous international agencies and utilized more than 120 GPS receivers from several different manufacturers with various antenna configurations. A data set was formed from the 21 JPL-developed Rogue GPS receivers [7]. These data were chosen to minimize the effects of antenna phase-center offsets and systematic errors internal to receivers. Moreover, the Rogue receivers provide dual-band P-code pseudorange and carrier-phase data types and utilize a low-multipath antenna. These features were sufficient for robust automatic editing of outliers and carrier-phase cycle repairs, thereby considerably simplifying the GPS analysis [8]. The Rogue receivers were globally distributed, as indicated in Fig. 1, where only the two fixed (fiducial) stations have been labeled explicitly. For a complete list of station names and coordinates, see [9]. Although the station network was global in scope, coverage was somewhat uneven, with 17 receivers in the Northern hemisphere but only four in the Southern hemisphere. Data were sampled continuously at a rate of one point every 2 minutes for most of the Rogue receivers. All the carrier-phase data were subsequently decimated to 6-min intervals, while the pseudorange data were smoothed by using the carrier phase to 6-min normal points.

The GPS data were reduced with the GPS Inferred Positioning System (GPSY) orbit-determination and baseline estimation software by using two basic strategies:

(1) a standard parameter-estimation strategy with two stations held fixed as fiducials [10], and (2) a variation of the above scheme with no fixed sites, i.e., a free-network strategy [9]. The standard strategy may be summarized as follows: Station locations, satellite states, and carrier-phase bias parameters were estimated as constants. Station and satellite clocks were estimated as white process noise, and the tropospheric delay for each station was estimated by using a random-walk stochastic model. When fiducial constraints were imposed, the station locations of Goldstone (California) and Kootwijk (Holland) were held fixed at coordinates taken from the SV5 reference frame [11]. In addition, the offset of the Earth's center of mass from the origin of the SV5 frame (the geocenter offset) was assumed to be zero.

The location of the rotation axis with respect to a crust-fixed axis (e.g., the IERS reference pole [12]) can be described by two coordinates, polar motion X (PMX) and Y (PMY), where the X-axis lies along the Greenwich meridian and the Y-axis is 90 degrees to the west (both orthogonal to the reference-pole Z-axis), as indicated in Fig. 1. The two pole parameters were estimated daily as constant adjustments to nominal values obtained from the IERS Bulletin B (B37 and B38).

All parameters were simultaneously estimated in a factorized Kalman filter. The satellite states were re-estimated daily to minimize systematic force model errors. The a priori sigmas of most estimated parameters were left essentially unconstrained. Solutions from separate, consecutive 24-hr data spans were computed; hence, pole positions were obtained daily as estimated offsets from the Bulletin-B values.

III. Results

A. Fiducial Versus Free-Network Strategies

Station locations were estimated as constants over the entire 3-week interval, defining a rigid polyhedron for which daily variations of the pole position could be estimated. Solutions were obtained from 22 separate 24-hr periods between January 22–30 and February 1–13. The daily GPS pole-position estimates for PMX and PMY versus time are shown in Fig. 2 and listed in Table 1. The GPS values are estimated corrections to the nominal Bulletin-B values, but the full PMX and PMY values are also shown in Table 1. The open symbols in Fig. 2 correspond to the fiducial case, with two stations fixed (in the SV5 frame). The GPS polar motion series exhibits both an apparent bias and periodic variability with respect to the IERS solution. The mean bias (GPS–Bulletin B) is 2.3 mas (7 cm)

in PMX and 6.7 mas (21 cm) in PMY. In addition to the biases, PMX and PMY exhibit variations of ~ 5 mas and ~ 2 mas, respectively, over 5–10 days. Note that the nominal IERS Bulletin-B reference values are themselves derived from smoothed points with 5-day spacing; hence, polar motion variability with a period shorter than ~ 15 days is not expected to be captured by the Bulletin-B values.

In the free-network case, the fiducial constraints were removed. All station locations were again estimated as constants over the 3-week period, but this time with large a priori sigmas for all stations. The filled triangles in Fig. 2 show the pole positions for this no-fiducial case after a mean offset has been removed. The rms agreement between the fiducial and free-network cases is ~ 0.03 mas (0.1 cm) for each pole component, which indicates that the fiducial strategy is unimportant for monitoring day-to-day pole-position variability (at least when a global, multi-day rigid network is available). In the free-network case, although the absolute orientation of the rigid network is ill-defined in the sense that it is not strongly tied to a terrestrial reference frame, the relative changes in the orientation of the network from day to day are apparently well determined. The absolute orientation of the rigid network is, however, weakly constrained by finite a priori sigmas (1 rad for PMX and PMY and 10 km for each station coordinate), and by fixed Bulletin-B values for UT1 (not solved for).

B. Consider Analysis

A number of sensitivity analyses [13] were run to determine the effects on pole-position estimates of potential errors in fiducial station coordinates and in geocenter offset. Fiducial errors, at the 1σ level, were assumed to be 3 cm in each of the three station coordinates and 10 cm for each of the geocenter components. These errors are of the same magnitude as the expected accuracies of the fiducial and geocenter positions within the SV5 terrestrial reference frame used to provide the fiducial station coordinates.

The effects on pole position due to these “consider” errors are illustrated in Fig. 3. Shown are the absolute values of the errors in both PMX and PMY averaged over seven 24-hr observing sessions. The root-sum-square (RSS) polar motion error due to errors in the three fiducial components at each site are shown with solid columns. Fiducial location errors on the order of 3–5 cm can apparently lead to pole-position errors at the 1-mas (3-cm) level. The crosshatched columns in Fig. 3 show the error in pole position due to the assumed error in each geocenter component. A 10-cm error in the geocenter can significantly

affect pole-position estimates, as errors in the X- and Y-axis geocenter components generate 2–3 mas pole-position errors. Note that the Z-axis geocenter error has a much weaker effect upon pole errors, which may be understood from geometrical arguments [14]. As seen in Fig. 2, fiducial (geocenter and station-coordinate) errors manifest themselves primarily as constant biases; therefore, they affect the absolute pole position but leave the day-to-day polar variations unaffected.

C. GPS Versus VLBI and SLR Estimates of Pole Position

Solutions from eleven daily VLBI experiments spanning January 23 to February 12 have been obtained [15]. Each solution was obtained from a 24-hr measurement on one of several VLBI networks utilizing typically 4–5 stations. To remove unevenness in network geometries, the station coordinates were fixed at values determined from approximately 1,300 experiments. A concise description of the analysis is given in [15]. Two pairs of solutions were centered only minutes apart; weighted averages were computed for these points so that effectively only nine VLBI pole-position estimates were available for comparison. The VLBI solution epochs occurred at varying times for each daily solution, whereas the GPS epochs occurred at 12:00 UTC each day (obtained from 24-hr solutions running from midnight to midnight). In order to compute the rms differences at the same epoch, the GPS pole-position estimates were linearly interpolated to the times of the VLBI measurements. The GPS and VLBI estimates are shown in Fig. 4, where a mean bias with respect to GPS has been removed from the VLBI solutions for each pole parameter to show the close correlation between the two polar motion series. The resulting rms differences between GPS and VLBI, after removing the mean biases, are 0.4 mas (1.2 cm) for PMX and 0.5 mas (1.5 cm) for PMY. The formal errors for the GPS pole-position estimates are typically 0.3–0.4 mas, while the VLBI formal errors range from 0.1–0.4 mas. Thus, the GPS and VLBI solutions are consistent to within 2σ . The biases between the GPS and VLBI solutions (for the GPS fiducial case, which utilizes the SV5 frame) are -0.1 mas in PMX and -3.2 mas for PMY, where the mean of the (VLBI–GPS) differences has been computed from nine points. Again, these offsets between VLBI and GPS are due mainly to reference frame differences and lie within the error bounds obtained from the consider analysis.

SLR data were obtained from the CSR 91 L 02 series produced by the Center for Space Research, University of Texas at Austin [1,2]. Each data point represents the mean of approximately three days’ worth of SLR measure-

ments. The SLR estimates are shown in Fig. 4 together with the GPS and VLBI results, where mean biases with respect to the GPS values have been removed from the SLR solutions to show the close correlation of all three polar motion series. The rms differences between GPS and SLR, after removal of the biases, are 0.3 mas (0.9 cm) for PMX and 0.4 mas (1.2 cm) for PMY. The formal errors for the SLR data average ~ 0.8 mas; therefore, the computed polar motion series from all three space geodetic techniques are consistent with one another. The mean biases for the (SLR-GPS) differences, computed from eight points, are -2.4 mas for PMX and -5.1 mas for PMY, due most probably to reference-frame inconsistencies.

IV. Conclusions

Day-to-day pole-position estimates spanning 3 weeks have been obtained with GPS by using data from a 21-station global tracking network. The GPS, VLBI, and SLR estimates agree within their formal errors after removing mean offsets, with typical RMS scatter of 0.4 mas (1.2 cm).

GPS fiducial errors do not appear to affect estimates of pole-position variations, but could introduce pole-position biases on the order of 2–3 mas (6–9 cm). Hence, GPS appears to yield polar motion estimates that are at least as precise as those from other techniques, with comparable accuracy as long as the relative biases are accounted for.

Just as significant, the GPS data provide polar motion estimates every 24 hours, and even higher time resolution is emerging from these data [16]. VLBI polar motion is routinely available with a typical spacing of 4–5 days, due to time and fiscal restrictions on the radio telescopes used by the various VLBI networks, while SLR data are only available every 3 days, due to the need to acquire several days worth of data to generate a reliable SLR normal point. GPS techniques do not suffer from either of these restrictions. Thus, GPS can be a powerful complementary technique to VLBI and SLR for monitoring high-frequency pole-position, which can lead to an improved DSN capability for tracking and navigation of future space missions with stringent tracking requirements.

Acknowledgments

The authors wish to thank Tom Herring for providing the VLBI solutions and Richard Eanes for the SLR data, as well as all the JPL research staff who were involved in processing the GIG'91 data. Discussions with R. Gross, A. Steppe, S. Lichten and T. Yunck were much appreciated.

References

- [1] IERS, *International Earth Rotation Service (IERS) Technical Note 5, Earth Orientation and Reference Frame Determinations, Atmospheric Excitation Functions, Up to 1989 (Annex to the IERS Annual Report for 1989)*, Paris: Central Bureau of IERS—Observatoire de Paris, June 1990.
- [2] IERS, *International Earth Rotation Service (IERS) Annual Report for 1990*, Paris: Central Bureau of IERS—Observatoire de Paris, July 1991.
- [3] A. P. Freedman, "Determination of Earth Orientation Using the Global Positioning System," *TDA Progress Report 42-99*, vol. July–September 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 1–11, November 15, 1989.
- [4] A. P. Freedman, "Measuring Earth Orientation With the Global Positioning System," *Bulletin Géodésique*, vol. 65, no. 1, pp. 53–65, March 1991.

- [5] S. M. Lichten, "Demonstration of Precise Estimation of Polar Motion Parameters With the Global Positioning System: Initial Results," *TDA Progress Report 42-104*, vol. October–December 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 1–8, February 15, 1991.
- [6] A. P. Freedman, S. M. Lichten, W. G. Melbourne, and S. Pogorelc, "A GPS Campaign to Assess High-Frequency Earth Rotation Variability" (abstract), *Eos Trans., AGU*, vol. 71, no. 43, p. 1272, October 23, 1990.
- [7] J. B. Thomas, "Functional Description of Signal Processing in the Rogue GPS Receiver," JPL Publication 88-15, Jet Propulsion Laboratory, Pasadena, California, June 1, 1988.
- [8] G. Blewitt, "An Automatic Editing Algorithm for GPS Data," *Geophys. Res. Lett.*, vol. 17, no. 3, pp. 199–202, March 1990.
- [9] M. Heflin, W. Bertiger, G. Blewitt, A. Freedman, K. Hurst, S. Lichten, U. Lindqwister, Y. Vigue, F. Webb, T. Yunck, and J. Zumberge, "Global Geodesy Using GPS Without Fiducial Sites," *Geophys. Res. Lett.*, in press, 1991.
- [10] S. M. Lichten and J. S. Border, "Strategies for High-Precision Global Positioning System Orbit Determination," *J. Geophys. Res.*, vol. 92, no. B12, pp. 12,751–12,762, November 10, 1987.
- [11] M. H. Murray, R. W. King, and P. J. Morgan, "SV5: A Terrestrial Reference Frame for Monitoring Crustal Deformation with the Global Positioning System" (abstract), *Eos Trans., AGU*, vol. 71, no. 43, p. 1274, October 23, 1990.
- [12] D. D. McCarthy, *International Earth Rotation Service (IERS) Technical Note 3, IERS Standards (1989)*, Paris: Central Bureau of IERS—Observatoire de Paris, November 1989.
- [13] G. J. Bierman, *Factorization Methods for Discrete Sequential Estimation*, Orlando, Florida: Academic Press, 1977.
- [14] Y. Vigue, S. M. Lichten, and G. Blewitt, "Precise Determination of Geocentric Station Coordinates with the Global Positioning System" (abstract), *Eos Trans., AGU*, supplement, pp. 114, October 29, 1991.
- [15] T. A. Herring, D. Dong, and R. W. King, "Sub-Milliarcsecond Determination of Pole Position Using Global Positioning System Data," *Geophys. Res. Lett.*, vol. 18, no. 10, pp. 1893–1896, 1991.
- [16] S. M. Lichten, "Very High Precision GPS Estimates of Changes in UT1–UTC and Polar Motion with Sub-Daily and Sub-Hourly Time Resolution from the GIG'91 Experiment" (abstract), *Eos Trans., AGU*, supplement, p. 120, October 29, 1991.

Table 1. GPS pole-position estimates using 2 fixed stations.^a

Date ^b	PMX ^c	PMY ^c	Full PMX ^d	Full PMY ^d
1/22/91	0.2 ±0.3	5.8±0.4	-50.4	93.1
1/23/91	0.9 ±0.3	5.9±0.4	-52.8	94.2
1/24/91	2.0 ±0.3	6.5±0.4	-54.9	96.1
1/25/91	3.7 ±0.3	6.7±0.4	-56.5	97.6
1/26/91	4.7 ±0.3	6.5±0.4	-58.7	98.7
1/27/91	4.6 ±0.3	7.7±0.4	-62.2	101.4
1/28/91	4.9 ±0.3	7.8±0.4	-65.5	103.0
1/29/91	3.6 ±0.3	7.3±0.4	-70.3	104.1
1/30/91	2.9 ±0.2	7.7±0.3	-74.8	106.0
2/1/91	1.4 ±0.2	7.5±0.3	-84.0	108.9
2/2/91	0.8 ±0.2	6.3±0.3	-88.4	109.2
2/3/91	0.9 ±0.2	6.7±0.3	-91.8	110.9
2/4/91	0.3 ±0.2	6.4±0.4	-95.9	111.9
2/5/91	0.04±0.2	6.7±0.3	-99.7	113.7
2/6/91	1.1 ±0.2	6.3±0.3	-102.2	114.9
2/7/91	1.4 ±0.2	6.1±0.3	-105.3	116.3
2/8/91	1.8 ±0.3	6.1±0.3	-108.3	118.1
2/9/91	2.0 ±0.3	6.1±0.3	-111.3	119.9
2/10/91	2.8 ±0.3	6.3±0.3	-113.6	122.1
2/11/91	2.7 ±0.3	6.2±0.3	-116.8	124.2
2/12/91	3.9 ±0.3	6.4±0.4	-118.7	126.6
2/13/91	2.7 ±0.3	6.2±0.4	-122.9	128.9

^a Units are in milliarcseconds.

^b All epochs are at noon UTC.

^c GPS corrections to the IERS Bulletins B37 and B38 nominal values.

^d GPS corrections added to the Bulletin B values linearly interpolated to noon UTC.

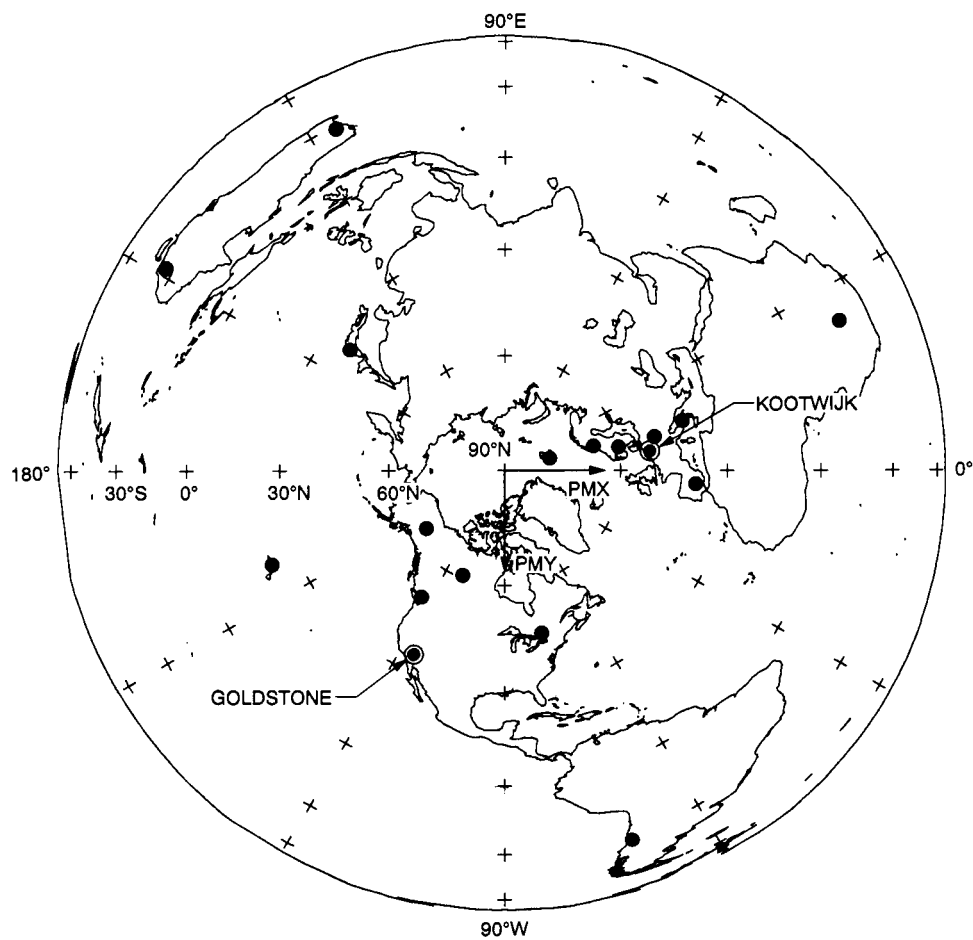


Fig. 1. Shown are 18 of the 21 Rogue sites, distributed worldwide during the GIG'91 experiment (3 additional Rogues in Southern California were also included in the analysis, with fiducial sites labeled).

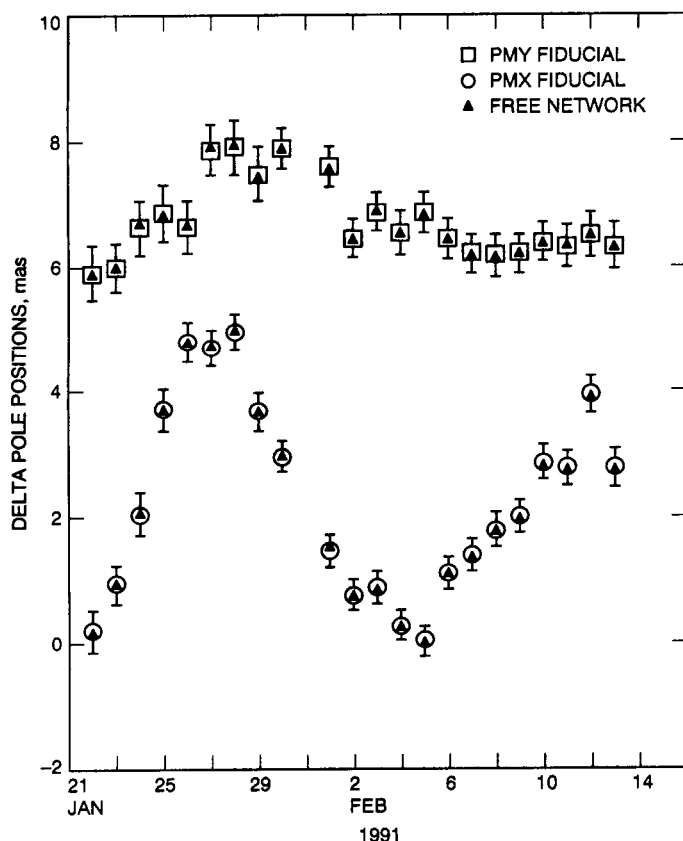


Fig. 2. A comparison of polar motion estimates for the fiducial versus the free-network cases.

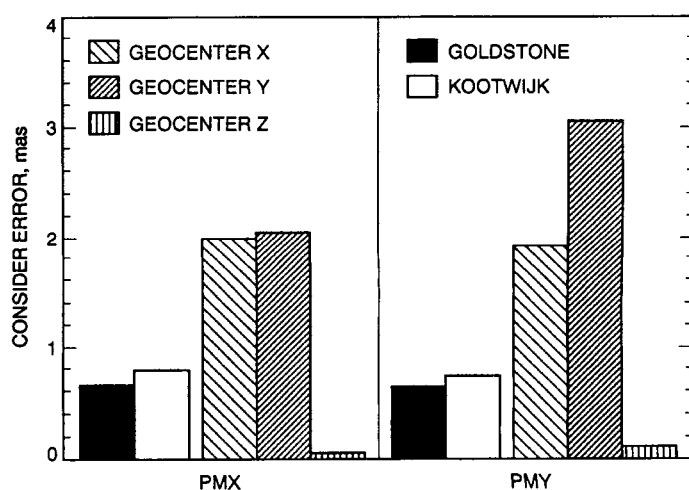


Fig. 3. Effect of consider errors on polar motion estimates. The solid bars show the consider errors based on a 3-cm error in each fiducial station coordinate, while the cross-hatched bars show the consider errors due to a 10-cm error in each geocenter component.

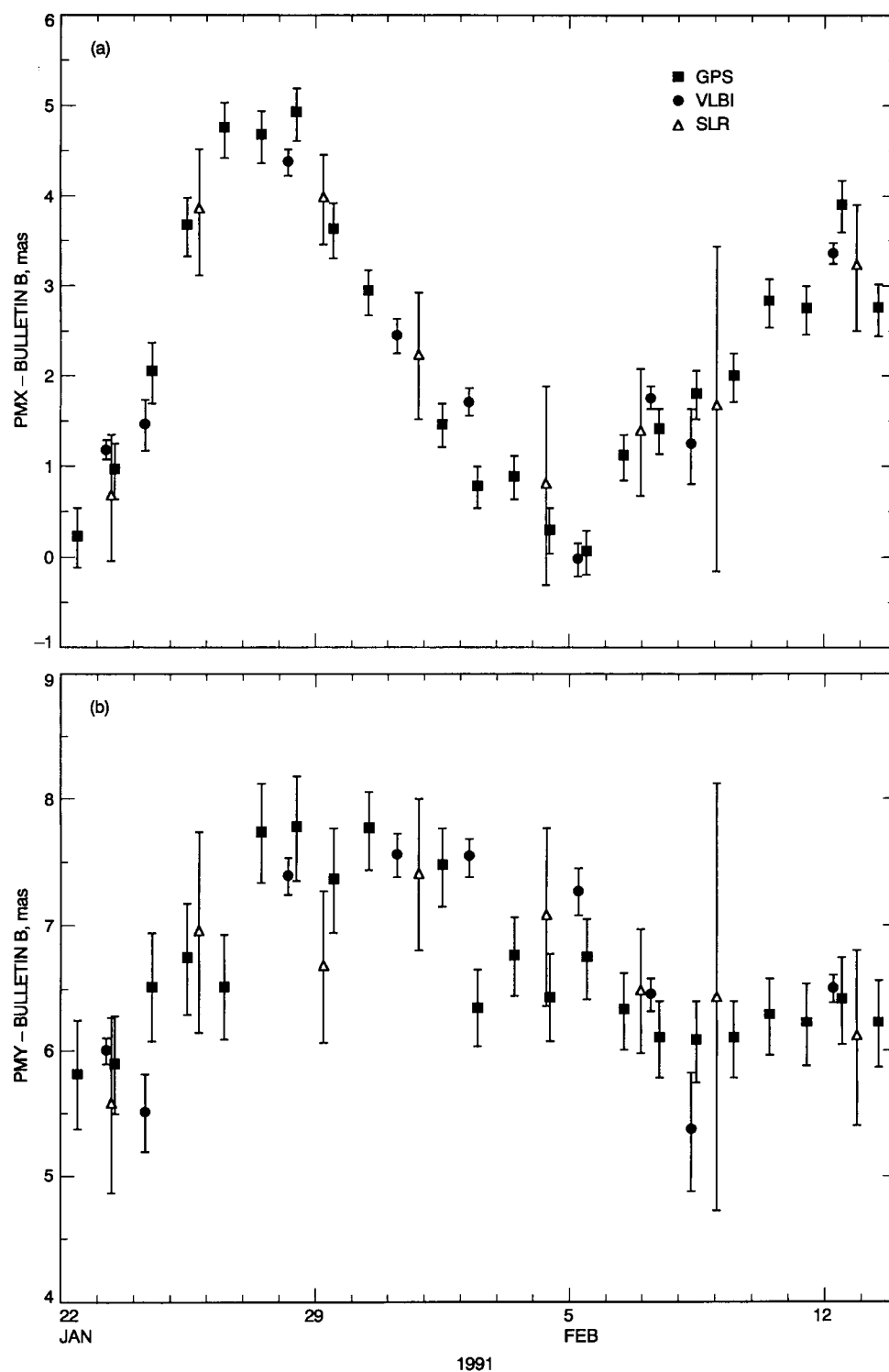


Fig. 4. Estimates of (a) polar motion X and (b) polar motion Y with GPS, VLBI, and SLR, with respect to Bulletin-B values. Mean differences with GPS have been removed from the VLBI and SLR data.

NO 2-24303

A High-Performance Hg^+ Trapped Ion Frequency Standard

J. D. Prestage, R. L. Tjoelker, G. J. Dick, and L. Maleki
Communications Systems Research

A high-performance frequency standard based on $^{199}\text{Hg}^+$ ions confined in a hybrid RF/dc linear ion trap has been demonstrated. This trap permits storage of large numbers of ions with reduced susceptibility to the second-order Doppler effect caused by the RF confining fields. A 160-mHz-wide atomic resonance line for the 40.5-GHz clock transition is used to steer the output of a 5-MHz crystal oscillator to obtain a stability of 2×10^{-15} for 24,000-second averaging times. Measurements with a 37-mHz line width for the Hg^+ clock transition demonstrate that the inherent stability for this frequency standard is better than 1×10^{-15} at 10,000-second averaging times.

I. Introduction

Atomic frequency standards with high stability for averaging times τ longer than 1000 seconds are necessary for a variety of astrophysical measurements and long baseline spacecraft ranging experiments. The millisecond pulsar, PSR 1937+21, shows stability in its rotational period that exceeds that of all man-made clocks for averaging times longer than 6 months. Comparison of this pulsar period with an Earth-based clock of stability 1×10^{-15} over averaging periods of one year may show the effects of very low frequency (≈ 1 cycle per year) gravitational waves [1,2]. Spacecraft ranging measurements across the solar system would be improved with clocks whose stabilities exceed 1×10^{-15} for averaging times of 10^4 to 10^5 seconds. This clock performance would also improve gravitational wave searches using spacecraft ranging data. Another use of clocks with long-term stability in NASA's Deep Space Network would be maintaining synchronization with Coordinated Universal Time (UTC).

stability ($\tau < 150$ seconds) the JPL Superconducting Cavity Maser Oscillator (SCMO) shows stability as good as 2×10^{-15} for averaging times extending to 2000 seconds [4].¹ Hydrogen masers are presently the most stable frequency standard for $150 < \tau < 30,000$ seconds and are the primary standards in use at each Deep Space Communications Complex. For averaging times greater than 6 months the millisecond pulsar PSR 1937+21 exceeds the stability of international timekeeping abilities at a level of 1 to 2 parts in 10^{-14} [3]. For $\tau > 10^6$ seconds the most stable clock yet measured is the Hg^+ ion standard based on 2×10^6 ions confined in an RF Paul trap [2,5]. The subject of the present article is a linear ion trap (LIT) based Hg^+ frequency standard now being tested at the Frequency Standards Laboratory (FSL) at JPL. Stability measurements of the Hg^+ standard using an H-maser local oscillator are shown in Fig. 1. Determination of the long-term stability ($\tau > 10,000$ seconds) is limited by instabilities in available reference H-masers.

The performance of microwave frequency standards in use today is summarized in Fig. 1 [2,3]. For short-term

¹ R. T. Wang, private communication, Jet Propulsion Laboratory, Pasadena, California, July 1991.

Typically the largest source of frequency offset in standards based on ions confined in Paul traps stems from the motion of the ions caused by the trapping fields via the second-order Doppler or relativistic time dilation effect. Though increasing ion number will lead to an increasing signal to noise ratio in the measurement of the clock transition, the frequency offset also grows with the number of ions forcing a trade-off situation. Often fewer ions are trapped in order to reduce the (relatively) large offset and frequency instabilities which may result.

II. Linear Ion Trap

In a conventional hyperbolic Paul trap, ions are trapped around a node of the RF electric field at the center. The strength of the electric field and the resulting micromotion of the trapped particles grows linearly with distance from this node point. As ions are added the size of the ion cloud grows until the second-order Doppler shift arising from the micromotion in the trapping field dominates the second-order Doppler shift from the ion's thermal motion at room temperature. For typical operating conditions [6] a spherical cloud containing 2×10^6 mercury ions shows a second-order Doppler shift of 2×10^{-12} , a value some 10 times larger than that due to thermal motion alone. In order to increase the number of stored ions with no corresponding increase in second-order Doppler shift from ion micromotion, a hybrid RF/dc LIT [7] was proposed² and developed. This trap confines ions along a line of nodes of the RF field effectively providing the same capability as a large number of hyperbolic traps together. The trapping force transverse to the line of nodes is generated by the ponderomotive force as in conventional Paul traps while the axial trapping force is provided by dc electric fields [7-10]. Unlike conventional RF traps this linear trap will hold positive or negative ions but not both simultaneously. A related trap using purely ponderomotive forces to confine charged particles is the racetrack trap [11,12].

The second-order Doppler shifts, $\Delta f/f$, generated by the trapping fields for a cloud of ions in these two types of traps can be compared, assuming that both traps are operated with the same RF trapping force as that characterized by the ion secular frequency ω . If the same number of ions N is held in both traps, the average distance from an ion to the node of the trapping field is greatly reduced in the linear trap. Since the distance from the node determines the magnitude of the RF trapping field, the second-order

Doppler shift of an ion's clock frequency due to motion in the trapping field is reduced from that of a hyperbolic trap. If R_{sph} is the ion cloud radius in the hyperbolic trap and L is the ion cloud length in the linear trap, the Doppler shift in the two traps is related by [7]

$$\left(\frac{\Delta f}{f}\right)_{lin} = \frac{5}{3} \frac{R_{sph}}{L} \left(\frac{\Delta f}{f}\right)_{sph} \quad (1)$$

As more ions are added to the linear trap this shift will increase. It will equal that of the spherical ion cloud in a hyperbolic trap when

$$N_{lin} = \frac{3}{5} \frac{L}{R_{sph}} N_{sph} \quad (2)$$

Equations (1) and (2) are valid when the ion cloud radii, R_{lin} and R_{sph} , are much larger than the Debye length, which is the characteristic plasma density decay length at the ion cloud edge and is about 0.4 mm for typical Hg^+ ion plasmas used in frequency standard work [6].

In addition to its larger ion storage capacity, the dependence of the second-order Doppler shift on trapping parameters in a linear trap is very different from that in a conventional Paul trap. For many ions in a Paul trap this shift is given by [6,7]

$$\left(\frac{\Delta f}{f}\right)_{sph} = -\frac{3}{10c^2} \left(\frac{N\omega q^2}{4\pi\epsilon_0 m}\right)^{2/3} \quad (3)$$

where ω is the secular frequency for a spherical ion cloud containing N ions each with charge to mass ratio q/m , c is the speed of light, and ϵ_0 is the permittivity of free space. Ions in a long linear trap where end effects are negligible show a second-order Doppler shift from the motion generated by the RF confining field given by [7]

$$\left(\frac{\Delta f}{f}\right)_{lin} = -\left(\frac{q^2}{8\pi\epsilon_0 mc^2}\right) \frac{N}{L} \quad (4)$$

where N/L is the linear number density of ions in the trap.

In contrast to the spherical case described in Eq. (3), this expression contains no dependence on trapping field strength, as characterized by ω , and depends only on the linear ion density N/L . If for example, the RF confining voltage increases and consequently the micromotion at a given point in space increases, the ion cloud radius will

² J. D. Prestage, *OSTDS Advanced Systems Review*, DSN RTOP 310-10-62 (internal document), Jet Propulsion Laboratory, Pasadena, California, pp. 64-68, June 14-15, 1988.

decrease so that the second-order Doppler shift (averaged over the cloud) from ion micromotion remains constant. Similar statements can be made about variations in any parameter that affects the radial confinement strength [8].

The sensitivity of the finite length linear trap to variations in radial trapping strength (characterized by ω) is [8]

$$\frac{\delta \left(\frac{\Delta f}{f} \right)_{lin}}{\left(\frac{\Delta f}{f} \right)_{lin}} = -2 \frac{R_t}{L} \frac{\delta \omega}{\omega} \quad (5)$$

and to variations in endcap voltage is

$$\frac{\delta \left(\frac{\Delta f}{f} \right)_{lin}}{\left(\frac{\Delta f}{f} \right)_{lin}} = 2 \cdot \frac{R_t}{L} \frac{\delta V_e}{V_e} \quad (6)$$

where R_t is the trap radius. The Paul trap shows a corresponding sensitivity to trapping field strength variations of

$$\frac{\delta \left(\frac{\Delta f}{f} \right)_{sph}}{\left(\frac{\Delta f}{f} \right)_{sph}} = -\frac{2}{3} \frac{\delta \omega}{\omega} \quad (7)$$

A comparison of Eqs. (5) and (7) shows the linear trap based frequency standard to be less sensitive to variations in trapping field strength than the Paul trap by a factor of $3R_t/L$. For the trap described in the next section this factor is about 1/3.

III. Operation With H-Maser Local Oscillator

Ions are produced inside the linear trap (Fig. 2) by an electron pulse along the trap axis which ionizes a neutral background vapor of ^{199}Hg . A helium buffer gas (10^{-5} mbar) collisionally cools the ions to near room temperature. Resonance radiation (194 nm) from a ^{202}Hg discharge lamp optically pumps the ions into the $F = 0$ hyperfine level of the ground state [13]. This UV light is focused onto the central 1/3 of the 75-mm-long ion cloud. The thermal motion of the ions along the length of the trap will carry all the ions through the light field so that pumping is complete in about 1.5 seconds for typical lamp intensities.

To minimize stray light entering the fluorescence collection system this state selection light is collected in a pyrex

horn as shown in Fig. 2. The location of the LaB_6 electron filament is also chosen to prevent light from entering the collection system. The filament's placement and relatively cool operating temperature, together with effective filtering of the state selection/interrogation UV light in the input optical system, allow frequency standard operation without the use of a 194-nm optical bandpass filter in the collection arm. This triples data collection rates since such filters typically have about 30-percent transmission for 194-nm light.

Microwave radiation (40.5 GHz), multiplied from an appropriate local oscillator (LO), passes through the trap perpendicular to the trap axis satisfying the Lamb-Dicke requirement that the spatial extent of the ion's motion along the direction of propagation of the microwave radiation be less than a wavelength. The microwaves enter through the pyrex horn (see Fig. 2) in opposition to the UV state selection/interrogation light. This allows collection of atomic fluorescence in both directions perpendicular to the incident pumping light. For the resonance and stability data shown in this article, fluorescence was collected in only one of these two directions.

The technique of successive oscillatory fields [14] has been used to probe the approximately 40.5-GHz hyperfine clock transition in $^{199}\text{Hg}^+$ ions confined to the linear trap. In the initial measurements the 40.5-GHz signal is derived from an active hydrogen maser frequency source as shown in Fig. 3. A representative resonance line of the $^{199}\text{Hg}^+$ clock transition is shown in Fig. 4. State selection and interrogation is accomplished during the 1.5 seconds following the lamp turn-on. It is necessary to switch the UV state selection/interrogation light level to near zero during the microwave interrogation period to prevent light shifts and broadening of the clock transition. A background light level of about 300,000 counts per 1.5-sec collection period has been subtracted to generate the resonance shown. The successive oscillatory field pulses consist of two 0.35-sec microwave pulses separated by a 2.5-sec free precession period. The data shown in Fig. 4 are averages of 10 scans with a 10-mHz frequency step size.

The central portion of the narrowest resonance lines yet obtained with this apparatus is shown in Fig. 5. This line is derived from two 0.275-second pulses separated by a 16.5-second free precession period [15]. The line width of 30 mHz represents a line Q of over 1×10^{12} on the 40.5-GHz transition. The data displayed are an average of four full scans and represent the highest Q transition ever measured in a microwave atomic transition.

To determine the frequency stability of the overall system of ions, trap, and microwave source, the output fre-

quency of the 40.5-GHz source has been locked to the central peak of the 160-mHz resonance in a sequence of 16,384 frequency measurements. The time required for each measurement is about 7 seconds and the loop response time was five measurement cycles. By averaging the frequencies of 2^N adjacent measurements ($N = 1, 2, \dots, 13$), the Allan deviation is formed as shown in Fig. 6(a). Performance of the linear ion trap based Hg^+ standard measured in this manner is $2 \times 10^{-13}/\sqrt{\tau}$ for averaging times $\tau \leq 20,000$ seconds, beyond which H-maser frequency instabilities limit the measurement. Figure 6(b) shows better performance obtained with a 37-mHz-wide resonance generated by two 0.35-second $\pi/2$ pulses with a 13-second precession period between. The stability measured in this mode of operation is 1×10^{-15} for an averaging time of 10,000 seconds.

At present no thermal regulation is incorporated into the ion standard itself. However, for all stability measurements described the ion trap standard together with its support electronics were housed in an environmentally controlled test chamber where temperature variations were regulated to approximately ± 0.05 deg C. Environmental sensitivities, leading perturbations, and control constraints of the Hg^+ trapped ion frequency standard are described and quantified in reference [19].

IV. Operation With a Crystal Oscillator

In this section is described the long-term stability achieved when a commercial (Oscilloquartz 8600B no. 104) 5-MHz quartz crystal oscillator of superior performance is servoed to the 160-mHz Hg^+ ion resonance as shown in Fig. 4. The schematic of this system is shown in Fig. 7. A frequency stability measurement of the Hg^+ steered crystal using two H-masers in the measurement system (designated Deep Space Network (DSN) 2 and DSN3) as references has been carried out using the JPL Frequency Standards Lab (FSL) measurement system.

The results of this 63-hour comparison are shown in Figs. 8(a) and (b). The first figure shows the Allan deviation of the Hg^+ steered crystal as measured against DSN2 (upper trace) and against DSN3 (lower trace). For averaging times less than 10 seconds the Allan deviation is that of the unsteered crystal oscillator since the loop attack time is about 10 seconds. For averaging times shorter than about 13,000 seconds the Hg^+ standard shows the same stability independent of a reference maser. For $\tau > 20,000$ seconds the Allan deviation of the Hg^+ versus DSN2 is the same as that for DSN2 versus DSN3 [see Fig. 8(b)] indicating that DSN2 has the limiting performance of the three standards

under test for this averaging time period. The lower trace of Fig. 8(a) shows a Hg^+ standard stability of 2×10^{-15} for $\tau = 24,000$ seconds beyond which instabilities in DSN3 probably limit the measurement. A second Hg^+ ion standard is now under construction, which will enable stability measurements beyond 24,000 seconds.

V. Local Oscillator Considerations

Fluctuations in the local oscillator (LO) limit performance of a trapped ion standard in two ways. As discussed above, slow variations of the crystal LO are compensated by action of the servo system. The effectiveness of this compensation increases with the measuring time, so that for longer measuring times performance approaches the $1/\sqrt{\tau}$ dependence which is characteristic of passive standards. This behavior is clearly shown in Fig. 8(a).

However, fast fluctuations in the LO also degrade performance of the standard by an effect that adds to the (white) fluctuation of the photon count from measurement to measurement [16–18]. This limitation continues to the longest times, having the same $1/\sqrt{\tau}$ dependence on measuring time τ as the inherent performance of the standard itself. The cause of this effect can be traced to a time varying sensitivity to LO frequency which is inherent in the interrogation process. This limitation was evaluated in a recent calculation for several types of sequentially interrogated passive standards [18].

Using this same methodology, one can calculate the LO-induced performance degradation for this particular interrogation scenario. Here, two RF pulses of 0.35-second length are separated by a delay of 2.5 seconds. A dead time of 3.8 seconds follows to give the cycle time of $t_c = 7$ seconds. The authors' quartz LO shows flicker frequency noise with an approximately flat Allan deviation of 1.5×10^{-13} . For this LO a contribution was calculated to the Allan deviation of the trapped ion source of $2.6 \times 10^{-13}/\sqrt{\tau}$. This value is slightly larger than the $2.0 \times 10^{-13}/\sqrt{\tau}$ due to random fluctuations in the photon count, again based on actual operating conditions. The two contributions can be combined to give a limiting stability of $3.3 \times 10^{-13}/\sqrt{\tau}$ which is plotted along with the data in Fig. 8(a). The data confirm this analysis by a very close approach to the line for measuring times $\tau > 10^3$ seconds.

Both LO and intrinsic statistical performance limitations may be reduced by increasing the interrogation time, as long as the dead time is not increased. The implicitly higher Q and reduced (relative) dead time would result in a comparable improvement for each of the two contribu-

tions. For example, an increase in the interrogation time by four times would reduce the limiting $1/\sqrt{\tau}$ Allan deviation by half.

However, it is clear that as performance improves to make possible trapped ion performance in the low $10^{-14}/\sqrt{\tau}$ range, quartz crystal LO performance will not be sufficient in itself to keep pace. In this case a cryogenic LO with 10^{-14} type stability for short times could be used, or a quartz LO with alternatively interrogated traps to give a uniform servo sensitivity with time [18].

VI. Conclusions

By steering a 5-MHz crystal oscillator to a 160-mHz atomic resonance ($Q = 3.3 \times 10^{11}$), performance of 2×10^{-15} has been measured for $\tau = 24,000$ seconds limited by the stability of available reference hydrogen masers. Line Q s as high as 1.3×10^{12} have been measured [15], indicating consequent performance for this trap as good as $5 \times 10^{-14}/\sqrt{\tau}$ for $\tau > 150$ seconds. While using a hydrogen maser as a local oscillator an Allan deviation of 1×10^{-15} at 10,000 seconds with a 37-mHz resonance line has been measured.

Acknowledgments

The authors thank R. Berends for making the ^{202}Hg lamps and W. Diener, A. Kirk and R. Taylor for assistance with the Frequency Standards Laboratory Test and Measurement System.

References

- [1] M. M. Davis, J. H. Taylor, J. M. Weisberg, and D. C. Backer, "High-Precision Timing Observations of the Millisecond Pulsar PSR1937+21," *Nature*, vol. 315, pp. 547-550, June 1985.
- [2] D. W. Allan, M. A. Weiss, and T. K. Peppler, "In Search of the Best Clock," *IEEE Trans. Instrum. Meas.*, vol. 38, no. 2, pp. 624-630, April 1989.
- [3] L. A. Rawley, J. H. Taylor, M. M. Davis, and D. W. Allan, "Millisecond Pulsar PSR 1937+21: A Highly Stable Clock," *Science*, vol. 238, pp. 761-765, November 1987.
- [4] R. T. Wang and G. J. Dick, "Improved Performance of the Superconducting Cavity Maser at Short Measuring Times," *Proceedings of the 44th Ann. Symp. on Freq. Control*, IEEE (IEEE Cat. no. 87-654207), Piscataway, New Jersey pp. 89-93, 1990.
- [5] L. S. Cutler, R. P. Giffard, P. J. Wheeler, and G. M. R. Winkler, "Initial Operational Experience with a Mercury Ion Storage Frequency Standard," *Proceedings of the 41st Ann. Symp. Freq. Control*, IEEE (IEEE Cat no. 87CH2427-3), Piscataway, New Jersey pp. 12-19, 1987.
- [6] L. S. Cutler, R. P. Giffard, and M. D. McGuire, "Thermalization of ^{199}Hg Ion Macromotion by a Light Background Gas in an RF Quadrupole Trap," *Appl. Phys. B*, vol. 36, pp. 137-142, January 1985.
- [7] J. D. Prestage, G. J. Dick, and L. Maleki, "New Ion Trap for Frequency Standard Applications," *J. Appl. Phys.*, vol. 66, no. 3, pp. 1013-1017, August 1989.

- [8] J. D. Prestage, G. R. Janik, G. J. Dick, and L. Maleki, "Linear Ion Trap for Second-Order Doppler Shift Reduction in Frequency Standard Applications," *IEEE Trans. Ultrason. Ferroelec. Freq. Contr.*, vol. 37, pp. 535-542, November 1990.
- [9] G. R. Janik, J. D. Prestage, and L. Maleki, "Simple Analytic Potentials for Linear Ion Traps," *J. Appl. Phys.*, vol. 67, no. 10, pp. 6050-6055, May 1990.
- [10] D. J. Wineland, J. C. Bergquist, J. J. Bollinger, W. M. Itano, D. J. Heinzen, S. L. Gilbert, C. H. Manney, and M. G. Raizen, "Progress at NIST Toward Absolute Frequency Standards Using Stored Ions," *IEEE Trans. Ultrason. Ferroelec. Freq. Contr.*, vol. 37, pp. 515-523, November 1990.
- [11] H. G. Dehmelt, "Introduction to the Session on Trapped Ions," *Proceedings of the 4th Symp. Frequency Standards and Metrology*, New York: Springer-Verlag, p. 286, 1989.
- [12] D. A. Church, "Storage Ring Ion Trap Derived from the Linear Quadrupole Radio Frequency Mass Filter," *J. Appl. Phys.*, vol. 40, no. 8, pp. 3127-3134, July 1969.
- [13] F. G. Major and G. Werth, "Magnetic Hyperfine Spectrum of Isolated $^{199}\text{Hg}^+$ Ions," *Appl. Phys.*, vol. 15, pp. 201-208, January 1978.
- [14] N. F. Ramsey, *Molecular Beams*, Oxford, England: Oxford University Press, 1956.
- [15] J. D. Prestage, G. J. Dick, and L. Maleki, "Linear Ion Trap Based Atomic Frequency Standard," *IEEE Trans. Instr. Meas.*, vol. 40, pp. 132-136, April 1991.
- [16] G. J. Dick, "Calculation of Trapped Ion Local Oscillator Requirements," *Proceedings of the 19th Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, Naval Observatory, Washington, D.C., pp. 133-146, 1988.
- [17] C. Audoin, V. Candelier, and N. Dimarcq, "A Limit to the Frequency Stability of Passive Frequency Standards," *IEEE Trans. Instr. Meas.*, vol. 40, pp. 121-125, April 1991.
- [18] G. J. Dick, J. D. Prestage, C. A. Greenhall, and L. Maleki, "Local Oscillator Induced Degradation of Medium-Term Stability in Passive Atomic Frequency Standards," *Proceedings of the 22nd Annual Precise Time and Time Interval (PTTI) Applications and Planning Meeting*, Naval Observatory, Washington, D.C., pp. 487-508, 1990.
- [19] R. L. Tjoelker, J. D. Prestage, and L. Maleki, "Criteria for Ultra-Stable Operation of the Trapped Ion Frequency Standard," *TDA Progress Report 42-108*, vol. October-December 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 19-30, February 15, 1992.

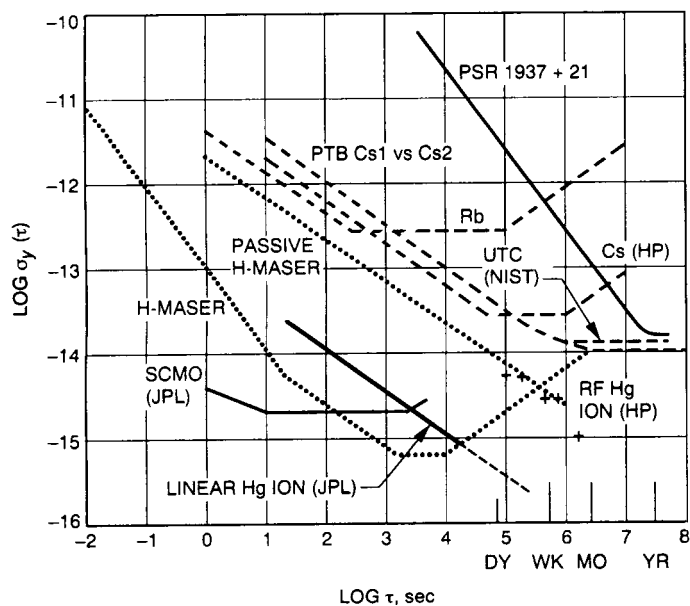


Fig. 1. Fractional frequency stability of several precision frequency standards. (PTB = Physikalisch-Technische Bundesanstalt; NIST = National Institute of Standards and Technology; HP = Hewlett Packard).

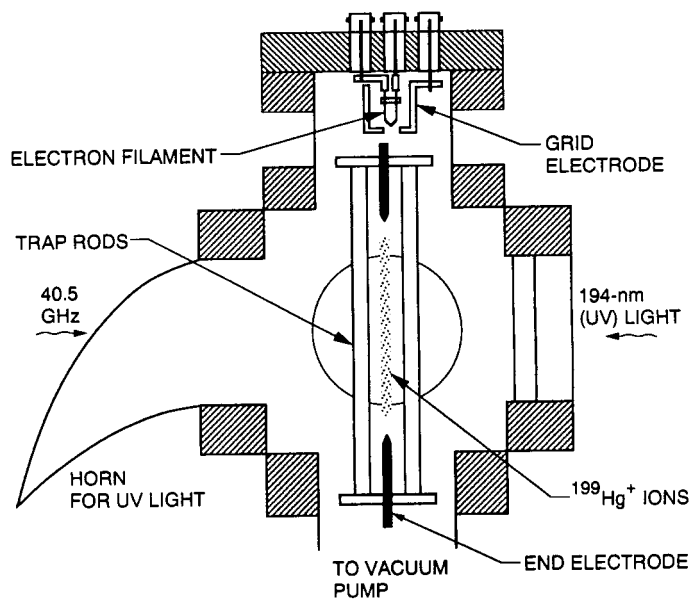


Fig. 2. Linear ion trap assembly residing in its high vacuum enclosure. State selection light from the ^{202}Hg discharge lamp enters from the right, is focused onto the central 1/3 of the trap and is collected in the horn. Fluorescence from the trapped ions is collected in a direction normal to the page.

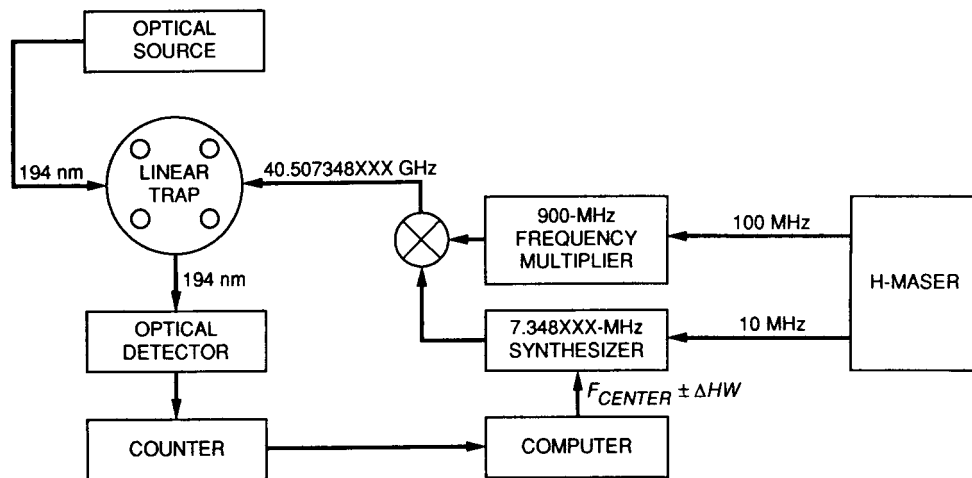


Fig. 3. Schematic operation of the mercury ion trap using a hydrogen maser as the local oscillator.

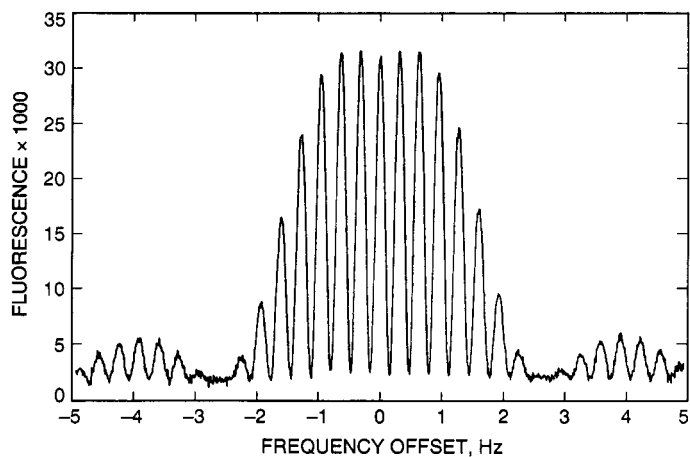


Fig. 4. $^{199}\text{Hg}^+$ clock transition as measured with method of successive oscillatory fields. This line shape results from two 0.35-second microwave pulses separated by a 2.5-second free precession period. The central line is about 160 mHz wide and the data shown are an average of ten scans.

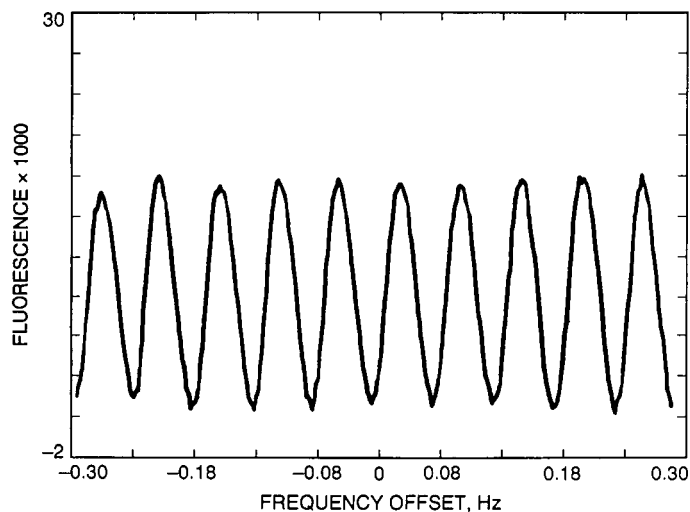


Fig. 5. High Q Ramsey fringes resulting from two microwave pulses of 0.275 seconds separated by a 16-second free precession period. The fringes are about 30 mHz wide and the data shown are an average of four scans.

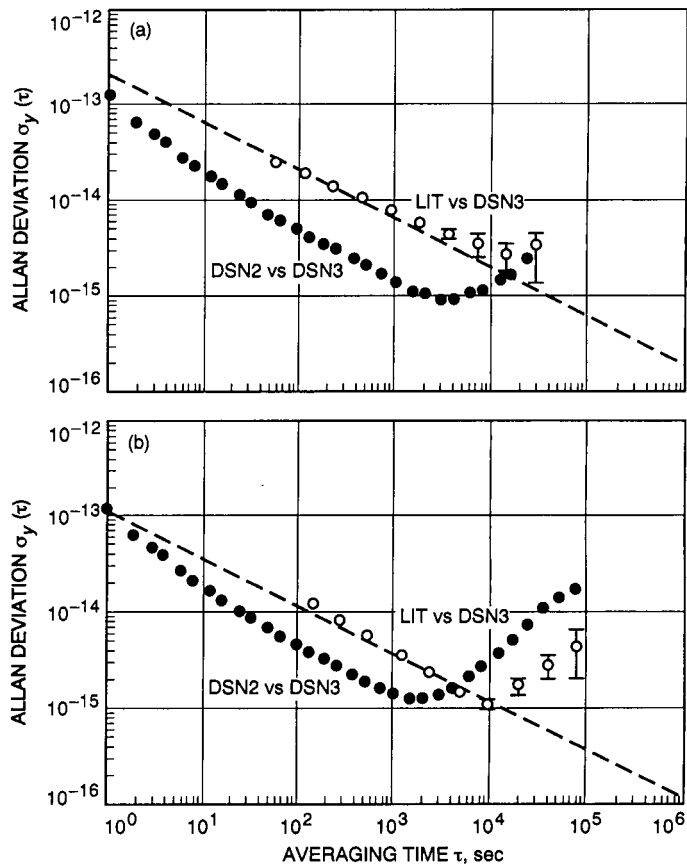


Fig. 6. Performance of the Hg^+ system when the local oscillator is a hydrogen maser; the fractional frequency stability of the ion trap system is measured against an LO maser: (a) Stability obtained with a 160-mHz Hg^+ resonance line as shown in Fig. 4; the $2 \times 10^{-13}/\sqrt{\tau}$ line is shown for reference. (b) Stability obtained with a 37-mHz Hg^+ resonance line. Also shown is the stability comparison of the LO maser to another H-maser.

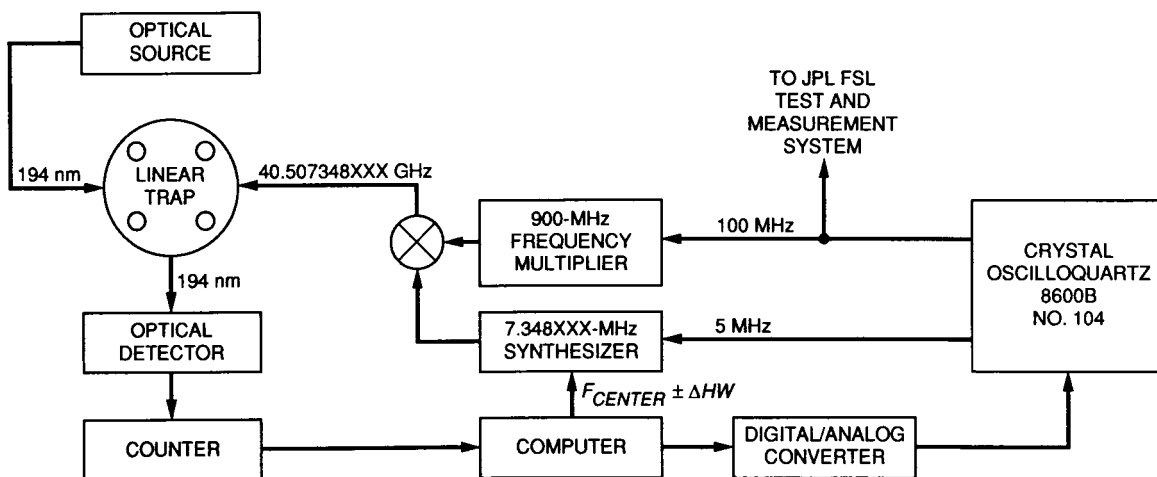


Fig. 7. Schematic operation of the mercury ion trap to steer a quartz oscillator.

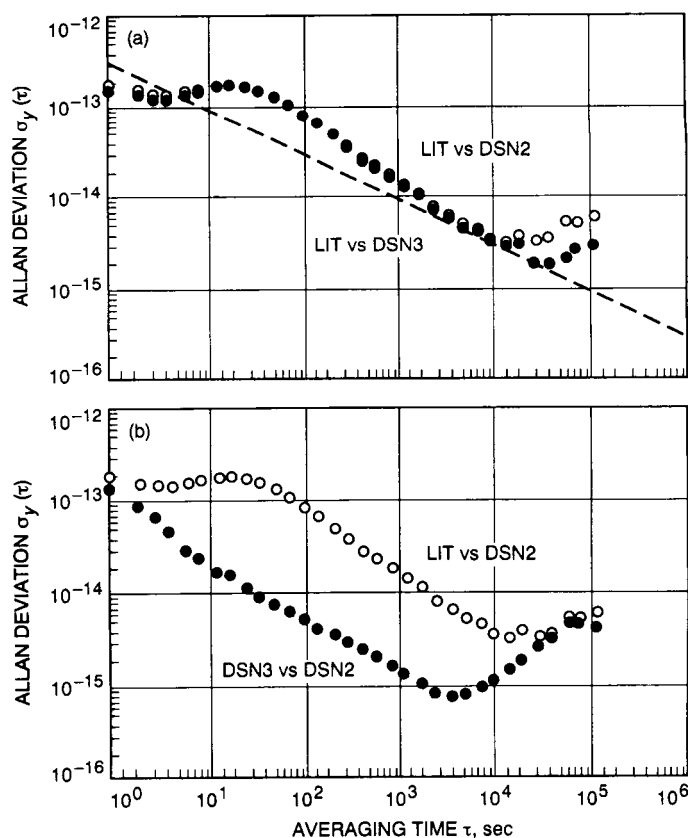


Fig. 8. Performance of the Hg^+ system when the LO is a quartz oscillator: (a) the ion trap standard compared to each maser independently, distinguishing the different performance of the two masers for times $\tau \geq 10,000$ seconds (the dashed line represents calculated $1/\sqrt{\tau}$ performance based on actual operating conditions), and (b) fractional frequency stability compared to the hydrogen maser DSN3 and maser comparison between DSN2 and DSN3.

53-32
79863
p 12
N92-24304
JJ574450

Criteria for Ultrastable Operation of the Trapped Ion Frequency Standard

R. L. Tjoelker, J. D. Prestage, and L. Maleki
Communications Systems Research Section

The leading systematic perturbations to the JPL mercury trapped ion frequency standard under present operating conditions are characterized. Sensitivity of the standard to environmental variations is measured, and the required regulation of key components to obtain a stability of 10^{-16} is identified.

I. Introduction

The 40.5-GHz atomic ground-state hyperfine "clock" transition of $^{199}\text{Hg}^+$ [1] has been measured in the JPL trapped ion frequency standard to have a corresponding line $Q \equiv \nu/\Delta\nu$ greater than 10^{12} , the highest ever measured in a microwave transition. A good signal to noise ratio (SNR) is achieved with as many as 3×10^7 ions in a linear ion trap (LIT)[5]. This high Q , together with improved SNR, has resulted in a demonstrated performance of $1 \times 10^{-13}/\sqrt{\tau}$ [2,3] where τ represents the averaging time interval.

The achievement of long-term stability depends upon the sensitivity to, and degree of isolation from, external influences. Because of their large mass and large hyperfine splitting, mercury ions provide a standard that, in general, is much less sensitive to environmental variations than hydrogen and cesium. Though sensitivities are greatly reduced, several offsets of the Hg^+ clock-transition frequency still exist under practical operating conditions. The obtainable long-term stability depends directly on the stabilization of these offsets. Under the present operating conditions, the fractional frequency stability of the trapped

ion standard has been measured to be better than 2×10^{-15} with 24,000-sec averaging times [2,3].

In this article, the leading perturbations and limitations to the present performance of the JPL mercury trapped ion frequency standard are quantified. Environmental sensitivity of the standard is measured, and the operating parameters required to reach a long-term stability of 10^{-16} are identified.

II. Operation of the Ion Trap Standard

In the present operational configuration (Figs. 1 and 2), $^{199}\text{Hg}^+$ ions are created inside the trap by an electron pulse along the trap axis, which ionizes a neutral background vapor of ^{199}Hg obtained by heating isotopically purified (≈ 96 -percent) HgO powder. Collisions with a helium buffer gas ($\approx 10^{-5}$ mbar) cool the ions to thermal equilibrium with the vacuum enclosure, currently kept at room temperature. Resonance ultraviolet radiation (UV) (194.2 nm) from a ^{202}Hg discharge lamp optically pumps the ions into the $F = 0$ hyperfine level of the ground state. To minimize scattered light, the state-selection light is collected in a pyrex horn. Microwave radiation (40.5 GHz)

to transfer the population to the $F = 1$ ground state enters the trap region through the pyrex horn and in a direction opposite to the UV state selection/interrogation light. This allows collection of atomic fluorescence in both directions perpendicular to the incident pumping light. Currently, fluorescence is collected in only one of these two directions. Inclusion of a second light-collection system should improve the SNR by $\sqrt{2}$.

The loading, state preparation, and interrogation cycle has been previously described [6]. After the $^{199}\text{Hg}^+$ ions are pumped into the $^2S_{1/2}(F = 0, m = 0)$ state, the technique of successive oscillatory fields [4] is used to probe the approximately 40.5-GHz hyperfine clock transition to the $^2S_{1/2}(F = 1, m = 0)$ state, Fig. 6(a). It is necessary to switch the UV state-selection/interrogation light level to near zero during the microwave interrogation period to prevent light shifts and broadening of the clock transition. After application of the second microwave pulse, the UV light is again turned back on to determine if the microwaves were on resonance and if they successfully transferred the population to the $F = 1$ state. A background light level between 10^5 and 5×10^5 counts per 1.0-sec collection period is subtracted to observe the resonance.

III. System Sensitivity and Systematics

A. Second-Order Doppler Shift

The largest frequency perturbation in standards based on ions confined in rf ion traps typically results from the motion of the ions (micromotion) due to the oscillating trapping fields via the second-order Doppler effect (relativistic time dilation). In order to increase the number of stored ions with little corresponding increase in second-order Doppler shift, a hybrid rf/dc linear ion trap is used [5]. This trap confines ions along a nodal line of the rf field. The trapping force transverse to the nodal line is generated by the ponderomotive force as in conventional hyperbolic Paul traps, while the axial trapping force is provided by dc electric fields.

Since the distance from the nodal line determines the magnitude of the rf trapping field, the second-order Doppler shift is reduced from that of a hyperbolic trap for the same number of ions by a factor of $(5/3)R_{\text{sph}}/L$, where R_{sph} is the radius of a hyperbolic trap and L is the length of the linear ion trap. Ions in a long linear trap (where end effects are negligible) show a second-order Doppler shift of

$$\left(\frac{\Delta f}{f}\right)_{\text{lin}} = -\left(\frac{q^2}{8\pi\epsilon_0 mc^2}\right)\frac{N}{L} \quad (1)$$

where N/L is the linear number density of ions in the trap [5]. For mercury, this fractional frequency shift is

$$\left(\frac{\Delta f}{f}\right)_{\text{Hg}} = -(3.8 \times 10^{-21} \text{ meters})\frac{N}{L} \quad (2)$$

Within the linear trap, the absolute second-order Doppler shift due to the trapping fields is greatly reduced, while allowing for an increased ion number and SNR.

Confinement lifetimes ranging from 2000–14,500 sec have been measured, depending upon the trap's well depth and background pressure. Even though the second-order Doppler shift and the sensitivity to ion-number fluctuations is reduced by using the linear ion trap, for ultrastable operation it still is necessary to keep the number of ions (N) in the trap fixed. The number is kept nearly constant (better than 0.1 percent) through repeated loading of the trap, though currently there is no active number stabilization. Once each interrogation cycle [6] (ranging from ≈ 5 –30 sec), a filament is pulsed, which emits electrons to ionize the neutral mercury in the trapping region. The equilibrium ion number is attained when the loading rate and loss rate are equal. The stability of the ion number over very long times depends on the stability of the background pressure of neutral mercury, filament emission, electron energy, electron pulse duration and duty cycle, and trap potentials (rf and dc). For the data presented in [2,3], these parameters have been regulated to a level summarized in Table 1. When the ionizing electrons are turned off, the number-dependent second-order Doppler shift from ion motion is directly observed by measuring the clock frequency as the ion number N decays (Fig. 3). By using Eq. (1) the measured frequency offset of 1.5×10^{-12} corresponds to about 3×10^7 ions, with the endcap potential of $V_{\text{EC}} = 10$ V.

If the background mercury pressure and electron pulse duration are sufficiently large, some passive stabilization is apparently achieved by filling the trap to a saturation level for a given voltage on the end electrodes. These electrodes are thin needles with a diameter of 2 mm [5]. Depending on the geometric details, the endcap electrodes can provide a limiting radius beyond which ions are no longer confined [7]. Under present operating conditions, the number stability is better than 0.1 percent, based on the measured stability of 2×10^{-15} over the 24,000-sec averaging times [2,3].

In Fig. 4(a), the frequency shift as a function of the potential V_{EC} on the endcaps is shown. This shift corresponds to a change in the average ion cloud radius. Figure 4(b) shows the measured peak amplitude (observed flu-

orescence) of the clock resonance versus the same variation in endcap voltage. This suggests that the frequency offset in Fig. 4(a) is due in part to variations in Hg^+ number. Further analysis of the potentials may provide a higher degree of number control in a passive way—a useful feature for a practical frequency standard.

B. Background Pressure Shifts

A buffer gas of neutral helium is used to cool and keep the Hg^+ ions in near-thermal equilibrium with the room-temperature vacuum enclosure. Slowing and cooling the ions allows them to become bound in the trap and reduces the radial extent of a given ion cloud, thus reducing the second-order Doppler shift. However, collisions between the mercury and helium buffer gas perturb the mercury hyperfine transition. The measured fractional frequency offset corresponding to helium pressure for a fixed set of loading conditions is shown in Fig. 5.

This dependence was first discussed by Cutler et al. [8]. At low pressures the quantity of helium is not sufficient to cool the trapped ions into thermal equilibrium with the vacuum enclosure. Instead, the mercury temperature rises with reduced helium pressure, which produces an increased thermal second-order Doppler shift of $3.2 \times 10^{-12}/10^{-5}$ mbar. For helium pressures higher than 6×10^{-6} mbar, the sensitivity is reduced to $1.1 \times 10^{-13}/10^{-5}$ mbar. In this region, the mercury ions are in thermal equilibrium and the predominant frequency dependence is a hyperfine shift due to collisions [9]. Currently, the helium pressure is measured with an ionization gauge, and the measured level is used to servo the heater of a quartz helium leak to hold the pressure constant. To reduce sensitivity to pressure fluctuations, yet operate with as little helium as possible, the operating pressure is typically 8×10^{-6} mbar.

In the present vacuum configuration (Fig. 1), the trap region is differentially pumped to a 40 l/sec turbo pump and the HgO source cannot be isolated from the trap vacuum. The partial helium pressure of 8×10^{-6} mbar constitutes the largest component of the background gas. The partial pressure of neutral ^{199}Hg is about 10^{-10} mbar. The helium is responsible for a fractional frequency offset of about 10^{-13} , and the presence of mercury (and presumably oxygen) most likely also perturbs the absolute frequency, though to a lesser extent, by several orders of magnitude.

Recently other gas contaminants have been reported to contribute significant systematic shifts in microwave atomic frequency standards. At the National Institute of Standards and Technology (NIST), a pressure-related shift

of $\approx 10^{-13}$ has been reported in the Be^+ hyperfine clock transition due to a 10^{-10} -mbar background of methane [10]. Control of trace amounts of background gas may be important in reaching the highest performances in ion traps. The best alternative is total elimination through source identification or perhaps cryogenic vacuums. An improved vacuum resulting in an increased confinement lifetime may also alleviate limitations due to number fluctuations.

C. Magnetism

The Zeeman energy levels of the ground state of $^{199}\text{Hg}^+$ as a function of magnetic field are shown in Fig. 6(a). The clock transition $^2S_{1/2}(F=0, m=0)$ to $^2S_{1/2}(F=1, m=0)$ is field-independent to the first order. The second-order field dependence is

$$\nu(B) = 40.507347997 \text{ GHz} + 97B^2 \left[\frac{\text{Hz}}{\text{G}^2} \right] \quad (3)$$

The fractional sensitivity of this transition to magnetic field variations is more than 800 times less than that of hydrogen and nearly 20 times less than with cesium at the same operating field. A Helmholtz coil centered on the trapping region is used to apply a constant magnetic field B . The minimum operating field depends upon the field homogeneity and interrogation time. Presently the operating field must be ≥ 35 mG. At 35 mG, the unshielded atomic sensitivity is $1.7 \times 10^{-13}/\text{mG}$. To reach 1×10^{-16} frequency stability, the current in the Helmholtz-field bias coils must be stable to 2×10^{-5} . To prevent ambient field disturbances from influencing the ion hyperfine frequency, the trap region is surrounded by a triple-layer magnetic shield with a shielding factor specified to be $\geq 10,000$. With this shielding factor, changes as large as 5 mG would lead to a 1×10^{-16} fractional frequency shift.

Fluctuations in the ambient magnetic field are typically less than 1 mG peak to peak. It is useful to examine these fluctuations over time and infer an equivalent Allan deviation (i.e., fractional frequency) to the magnetic environment. Figure 7(a) shows measurements of magnetic fluctuations averaged over time and scaled with the trapped ion sensitivity (for $B = 50$ mG and a shielding factor of 10,000). This trace exemplifies the inherent insensitivity of the mercury ion standard to magnetic field fluctuations. Currently, magnetic fluctuations are not a limiting systematic until below 10^{-16} stability.

Magnetic gradients are more of an immediate concern towards reaching higher frequency stability since gradients over the ion cloud will degrade atomic coherence and limit

line Q. The linear trap geometry is susceptible to field inhomogeneities, and field gradients are a limitation to long interrogation times. Even though the materials of the trap are carefully selected for their nonmagnetic properties, small paramagnetism of the trap, vacuum enclosure, and supporting materials result in small gradients across the ion cloud. On the same order, gradients also result from the field produced by the Helmholtz coils that are located inside the magnetic shields. To achieve high line Q, a partial solution has been to operate the standard in a relatively high magnetic field (≥ 35 mG).

A related situation resulting in loss of line Q is shown in Figs. 6(b) and 6(c). In this case, the Ramsey signal is lost when the Zeeman splitting of the magnetic states $^2S_{1/2}(F=1, m=\pm 1)$ is near the ion motional frequency of 500 kHz. This region must be avoided by proper choice of the trapping frequency in relation to the operating magnetic field.

D. Optical Pumping and Detection System

The optical pumping for state selection and interrogation with 194-nm light is a most critical area, one where improvements may lead to substantial gains in performance through increased SNR. The greatest difficulty is minimizing unwanted scattered light due to the use of a lamp as the optical pumping light source. Laser excitation would greatly improve the SNR through low background rates. Unfortunately, laser systems at 194 nm are not presently suitable for practical frequency standards that perform with long-term stability. With the present optics system, the achieved SNR is typically 40.

The lamp is dimmed during the Ramsey interrogation, since a light shift of $1.2 \times 10^{-15}/1000$ counts is observed. The typical counts during each interrogation result in an offset of 10^{-15} , though fluctuations in the count number are an order of magnitude lower.

E. Microwave Source: Local Oscillator

The multiplier chain (100 MHz–40.5 GHz), amplifiers, cabling, and waveguide to generate and transport the approximately 40.5-GHz microwave signal to drive the Hg^+ transition are not thermally regulated in the current apparatus. Because of the high degree of temperature control in the Frequency Standards Laboratory (FSL), this has not been a limitation at JPL. The microwave source did show sensitivity to temperature fluctuations during recent tests at Goldstone station DSS 13. An engineering version will need thermal regulation.

The short-term performance of the passive ion trap standard is dependent on the local oscillator (LO) and

multipliers used to produce 40.5 GHz. Hydrogen masers, excellent quartz crystals [2,3], and the JPL superconducting cavity maser oscillator (SCMO) have been used as the LO. In a free-run mode, stability is limited by the LO, and for practical purposes this mode of operation is only valuable for diagnostics and systematic measurements. With feedback, all three of these sources have been steered, and in principle compensate for environmental shifts in the LO.

F. Temperature Shifts

At the present time, no thermal regulation is incorporated into the ion standard itself, and only the helium-leak heater is actively regulated. The fractional temperature sensitivity of the complete system is less than $10^{-14}/\text{deg C}$ as measured by varying the temperature in one of the test chambers at JPL's FSL. Using this sensitivity, the temperature control in the test chamber is displayed in Fig. 7(b) as an inferred Allan deviation. Most of the measured sensitivity seems due to an increase in neutral mercury vapor and a consequent increase in ion number (and therefore cloud radius and second-order Doppler shift). The mercury lamp and its housing must also be temperature controlled because the 194-nm UV brightness is highly dependent on the operating temperature. In the present version, the temperature and local atmosphere are controlled by blowing dry nitrogen gas over the tip of the bulb, which controls the ^{202}Hg pressure and prevents ozone formation around the lamp.

Since the lamp (on average ≤ 7 W) is inside the magnetic shields, the equilibrium-operating temperature depends on the duty cycle of the lamp operation. The time constant for the lamp and surrounding apparatus to reach thermal equilibrium is about 3 hr after starting with new settings. A recent addition of a nonmagnetic shutter constructed from a piezoelectric bimorph now allows the lamp to run continuously, which eliminates temperature fluctuations when changing to different duty cycles. Small temperature changes in the vacuum enclosure would also slightly raise the equilibrium helium temperature, though the second-order Doppler shift due to thermal motion is significantly lower than the shift due to the trapping fields (Fig. 8).

The neutral mercury pressure in the trap region depends on the temperature of the HgO source and on the details of the sublimation process, which may vary with time. Temperature control on the heater is currently about 0.2 deg C out of 200 deg C. Higher regulation is needed, though an active ion-number control would eliminate the need for high regulation on this heater and the reliance on the stability of the HgO sublimation process and rate.

G. Transportability and Vibration Sensitivity

In order to determine sensitivity to transportation and actual field conditions, the unit was transported to the Goldstone complex and operated at DSS 13. The "laboratory" unit successfully demonstrated transportability, and the complete system was operational less than 24 hr after arrival at Goldstone from a cold start. The overall performance was similar to that at the FSL. Since no thermal regulation was integrated into the standard, temperature fluctuations at DSS 13 were the limiting systematic.

The ion trap was under study in the FSL during the June 28, 1991 magnitude-5.8 earthquake. Table 2 shows the relative frequency shift between frequency standards under test at the FSL and to the NIST time scale via the Global Positioning System (GPS). The JPL hydrogen masers DSN2 and DSN3 showed large shifts, while the ion trap showed no shift with respect to the commercial standards by HP (Cesium clock) and SAO (Hydrogen Maser) as well as the NIST (GPS) reference. Unlike these commercial standards, the present "laboratory" version of the ion trap has no vibration isolation incorporated. The earthquake exemplifies the insensitivity of the trapped ion standard to mechanical stress and low-frequency vibration.

IV. Summary

Short-term performance can best be improved by increased SNR and higher line Q. Improved optical systems

and geometries offer the most promise for achieving significant improvements in SNR. Improved magnetic homogeneity and vacuum control are the most critical areas for achieving even higher line Q's.

The environmental sensitivity of the JPL linear ion trap frequency standard has been measured. The long-term stability depends on the sensitivity of the atomic system to changes in environmental and operating parameters and on the ability to control and stabilize such parameters. The major frequency offsets under current operating conditions are summarized in Fig. 8, and the system sensitivities, present regulation, and regulation required for a stability of 10^{-16} are summarized in Table 1. The ion trap serves as a practical high-performance frequency standard with long-term stability as good as the regulation will allow. Long-term regulation of the ion number appears to be the most immediate challenge in reaching higher frequency stabilities. In the present mode of operation, further temperature regulation of the HgO heater will help, though active ion-number stabilization may be required.

A second system now under development will entail a higher degree of regulation and engineering, which is expected to improve long-term performance. Since measurements greater than 24,000 sec are already limited by the reference hydrogen masers [2,3], this second system will also provide the means for measuring and quantifying the achievable long-term stability with ion trap technology.

Acknowledgments

The authors thank W. Diener, A. Kirk and R. Taylor for assistance with the Frequency Standards Laboratory (FSL) Test and Measurement System and R. Wang and C. Greenhall for acquisition and determination of the inferred Allan deviations for magnetic and temperature fluctuations at the FSL.

References

- [1] F. G. Major and G. Werth, "Magnetic Hyperfine Spectrum of Isolated $^{199}\text{Hg}^+$ Ions," *Appl. Phys.*, vol. 15, pp. 201-208, January 1981.
- [2] J. D. Prestage, R. L. Tjoelker, G. J. Dick, and L. Maleki, "Ultra-Stable Hg^+ Trapped Ion Frequency Standard," in special issue on trapped ions, *J. Mod. Opt.* (1992); in press.
- [3] J. D. Prestage, R. L. Tjoelker, G. J. Dick, and L. Maleki, "A High-Performance Hg^+ Trapped Ion Frequency Standard," *TDA Progress Report 42-108*, vol. October-December, Jet Propulsion Laboratory, Pasadena, California, pp. 10-18, February 15, 1992.
- [4] N. F. Ramsey, *Molecular Beams*, New York: Oxford University Press, 1956.
- [5] J. D. Prestage, G. J. Dick, and L. Maleki, "New Ion Trap for Frequency Standard Applications," *J. Appl. Phys.*, vol. 66, no. 3, pp. 1013-1017, August 1989.
- [6] J. D. Prestage, G. J. Dick, and L. Maleki, "The JPL Trapped Mercury Ion Frequency Standard," *TDA Progress Report 42-92*, vol. October-December, Jet Propulsion Laboratory, Pasadena, California, pp. 13-19, February 15, 1988.
- [7] R. K. Melbourne, J. D. Prestage, and L. Maleki, "Inhomogeneous Electric Field Effects in a Linear RF Quadrupole Trap," *TDA Progress Report 42-101*, vol. January-March, Jet Propulsion Laboratory, Pasadena, California, pp. 51-57, May 15, 1990.
- [8] L. S. Cutler, R. P. Giffard, and M. D. McGuire, "Thermalization of ^{199}Hg Ion Macromotion by a Light Background Gas in an RF Quadrupole Trap," *Appl. Phys.*, vol. B 36, pp. 137-142, March 1985.
- [9] J. Vetter, M. Stuke, and E. W. Weber, "Hyperfine Density Shifts of $^{137}\text{Ba}^+$ Ions in Noble Gas Buffers," *Z. Physik*, vol. A 273, no. 2, pp. 129-135, 1975.
- [10] J. J. Bollinger, D. J. Heinzen, W. M. Itano, S. L. Gilbert, and D. J. Wineland, "A 303-MHz Frequency Standard Based on Trapped Be^+ Ions," *IEEE Trans. Instr. Meas.*, vol. 40, no. 2, pp. 126-128, April 1991.

Table 1. Present sensitivity and regulation of the trapped Hg⁺ frequency standard.

Systematic	Sensitivity	Present regulation ^a	Required for 10 ⁻¹⁶ stability ^b
Temperature			
FSL chamber (23 deg C)	$\left(\frac{\Delta\nu}{\Delta T}\right) \frac{1}{\nu} = -9 \times 10^{-15} \left[\frac{1}{\text{deg C}}\right]$	$\Delta T \leq 0.08 \text{ deg C}$	$\Delta T \leq 0.01 \text{ deg C}$
Magnetic field			
Helmholtz supply ($B = 35 \text{ mG}$)	$\left(\frac{\Delta\nu}{\Delta B}\right) \frac{1}{\nu} = 1.7 \times 10^{-13} \left[\frac{1}{\text{mG}}\right]$	$\frac{\Delta I}{I} \leq 2 \times 10^{-6}$	$\frac{\Delta I}{I} \leq 2 \times 10^{-5}$
Ambient fluctuation ($\Delta B_{\text{out}} \leq 5 \text{ mG}$)		$\frac{\Delta B_{\text{in}}}{\Delta B_{\text{out}}} \approx 10^{-4}$	$\frac{\Delta B_{\text{in}}}{\Delta B_{\text{out}}} \leq 10^{-4}$
Pressure			
Helium ($P = 8 \times 10^{-6} \text{ mbar}$)	$\left(\frac{\Delta\nu}{\Delta P}\right) \frac{1}{\nu} = 1.1 \times 10^{-13} \left[\frac{1}{10^{-5} \text{ mbar}}\right]$	$\frac{\Delta P}{P} \leq 0.009$	$\frac{\Delta P}{P} \leq 0.001$
Second-order Doppler (ion-number dependence)			
Endcap potential ($V \geq 10 \text{ V}$)	$\left(\frac{\Delta\nu}{\Delta V}\right) \frac{1}{\nu} = -2.4 \times 10^{-15} \left[\frac{1}{V}\right]$	$\Delta V \leq 1 \text{ mV}$	$\Delta V \leq 1 \text{ mV}$
($V \leq 6 \text{ V}$)	$\left(\frac{\Delta\nu}{\Delta V}\right) \frac{1}{\nu} = -2.5 \times 10^{-13} \left[\frac{1}{V}\right]$	$\Delta V \leq 1 \text{ mV}$	$\Delta V \leq 0.3 \text{ mV}$
Filament emission ($300\mu\text{A}$)	$\left(\frac{\Delta\nu}{\Delta I}\right) \frac{1}{\nu} = -1.2 \times 10^{-15} \left[\frac{1}{\mu\text{A}}\right]$	$\frac{\Delta I}{I} \leq 3 \times 10^{-3}$	$\frac{\Delta I}{I} \leq 3 \times 10^{-4}$
HgO heater ($\approx 200 \text{ deg C}$)($V \geq 10 \text{ V}$)	$\left(\frac{\Delta\nu}{\Delta T}\right) \frac{1}{\nu} = -2 \times 10^{-14} \left[\frac{1}{\text{deg C}}\right]$	$\Delta T_{\text{HgO}} \leq 0.2 \text{ deg C}$	$\Delta T_{\text{HgO}} \leq 0.004 \text{ deg C}$
($V \leq 6 \text{ V}$)	$\left(\frac{\Delta\nu}{\Delta T}\right) \frac{1}{\nu} = -8 \times 10^{-15} \left[\frac{1}{\text{deg C}}\right]$	$\Delta T_{\text{HgO}} \leq 0.2 \text{ deg C}$	$\Delta T_{\text{HgO}} \leq 0.01 \text{ deg C}$
Light shift			
Dim lamp ($\approx 3000 \text{ Hz}$)	$\left(\frac{\Delta\nu}{\Delta \gamma}\right) \frac{1}{\nu} = 1.2 \times 10^{-16} \left[\frac{1}{100 \text{ Hz}}\right]$	$\frac{\Delta \gamma}{\gamma} \leq 0.07$	$\frac{\Delta \gamma}{\gamma} \leq 0.03$

^a Peak-to-peak fluctuation or drift over $2 \times 10^5 \text{ sec}$.

^b Worst-case assumes no averaging of fluctuations.

Table 2. Permanent frequency offset of Frequency Standards Laboratory clocks as compared with hydrogen masers DSN2 and DSN3, which resulted from the 5.8-magnitude earthquake of June 28, 1991.

DSN2 versus DSN3	4×10^{-11}
DSN3 versus HP 5061B (cesium)	$-2.05(5) \times 10^{-12}$
DSN3 versus SAO26 (H-maser)	$-2.05(5) \times 10^{-12}$
DSN3 versus JPL LIT (Hg ⁺ linear ion trap)	$-2.06(5) \times 10^{-12}$
DSN3 versus NIST/GPS	$-1.97(5) \times 10^{-12}$

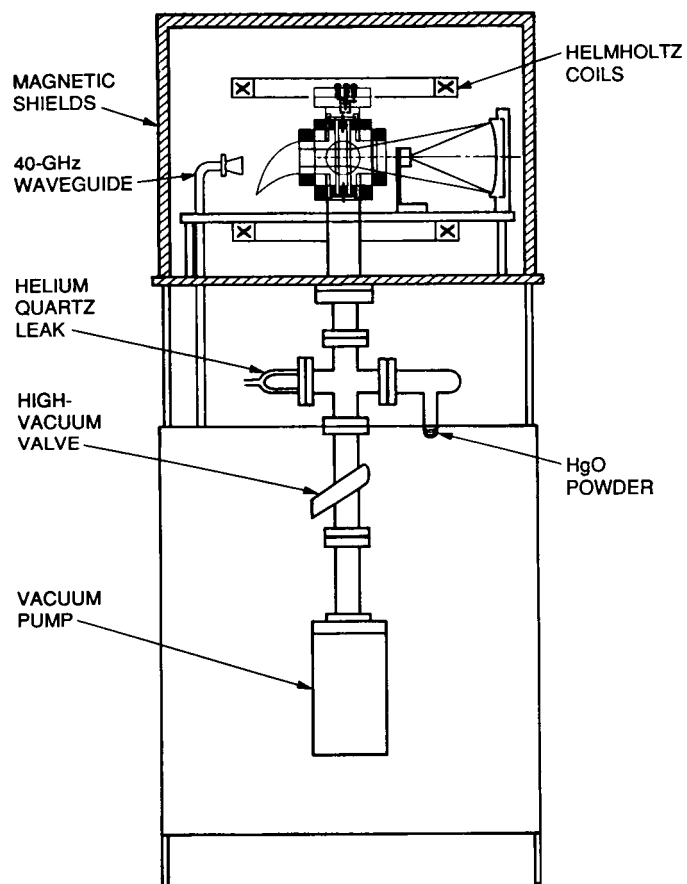


Fig. 1. Linear ion-trap frequency standard. State selection light from the ^{202}Hg discharge lamp enters from the right and is collected in a pyrex horn. Fluorescence from the trapped ions is collected in a direction normal to the page.

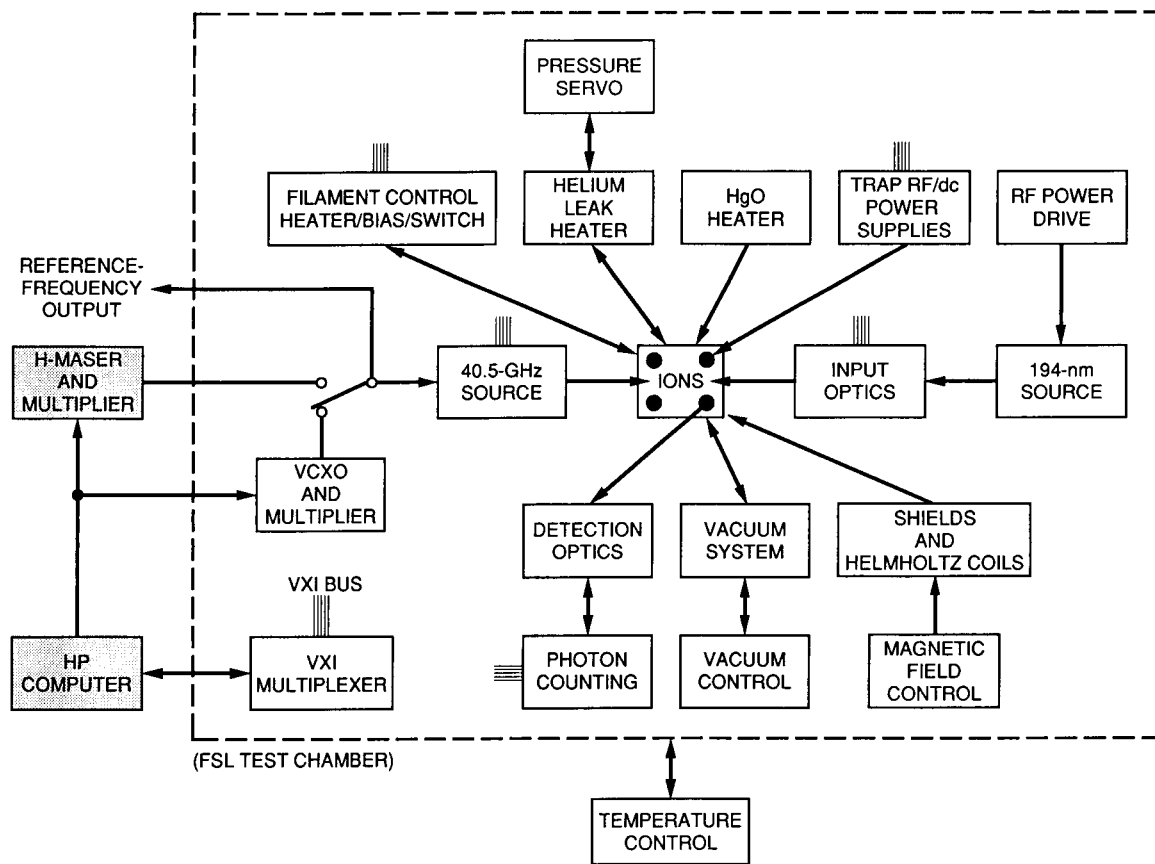


Fig. 2. The standard showing key components and subsystems.

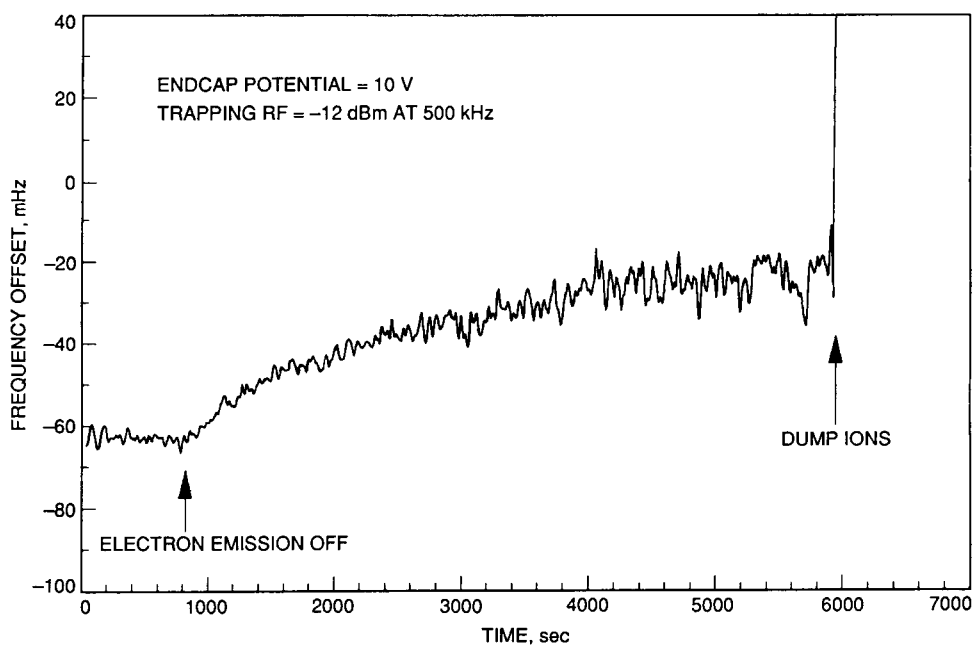


Fig. 3. A direct measure of the second-order Doppler shift due to ion motion in the rf trapping fields. As the trapped ion number diminishes, the clock frequency increases.

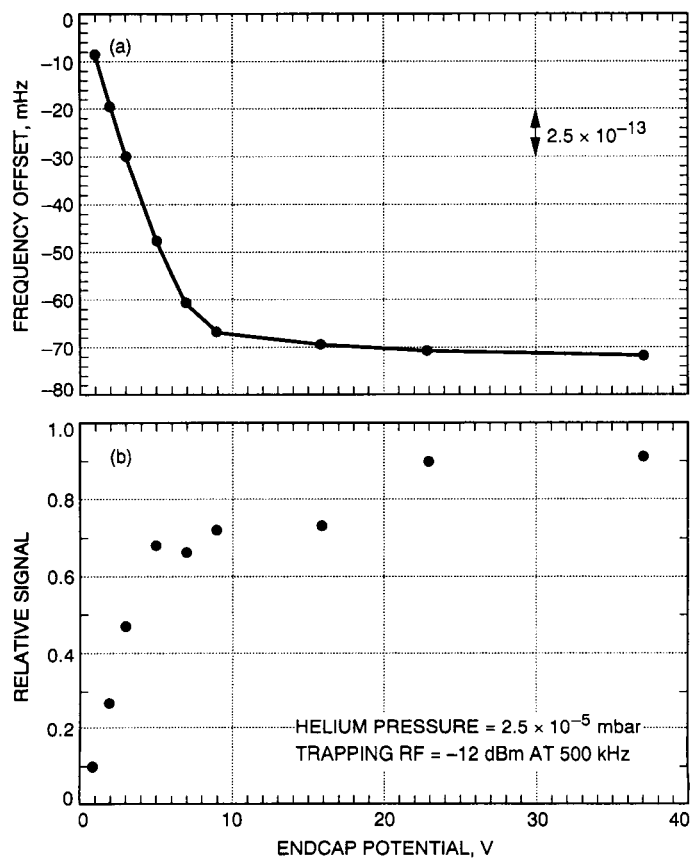


Fig. 4. Observed variations as a function of endcap potential: (a) fractional frequency offset of the clock transition and (b) signal peak amplitude.

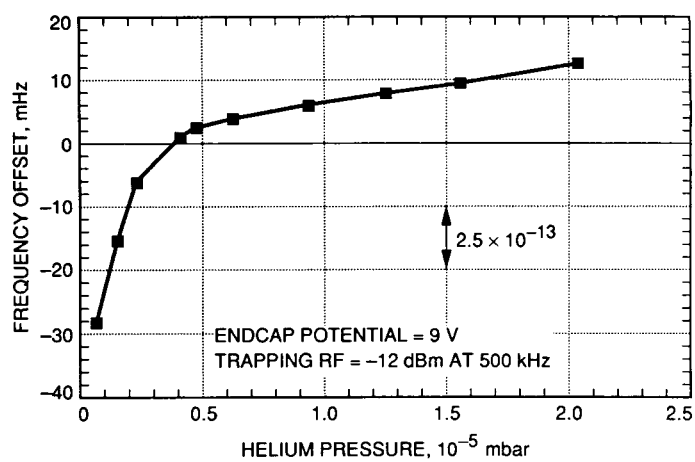


Fig. 5. The fractional frequency offset versus helium-buffer gas pressure in the trap region.

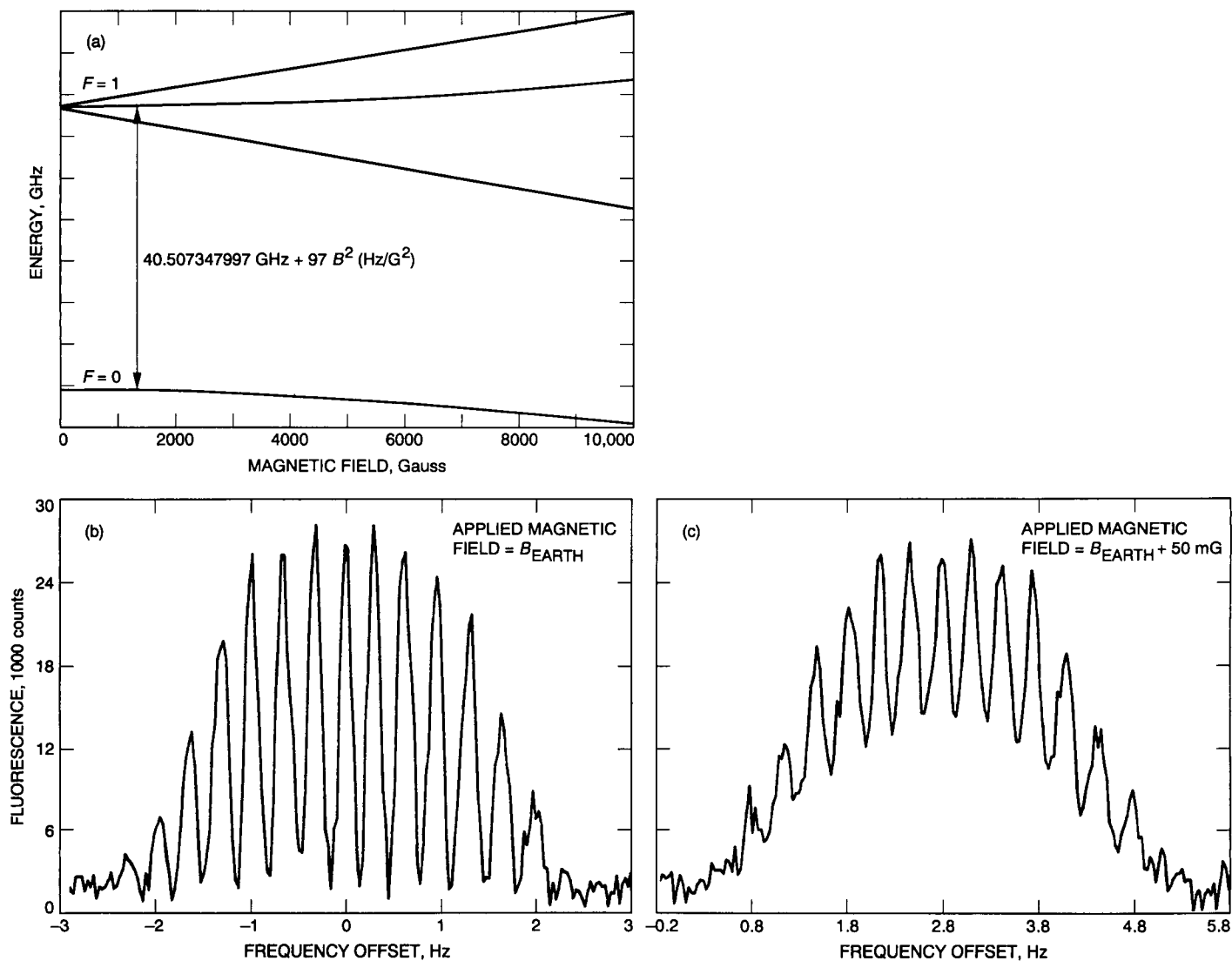


Fig. 6. Ground state transition in $^{199}\text{Hg}^+$: (a) schematic energy levels as a function of magnetic field; (b) typical line shape using the Ramsey method of successive oscillatory fields, with coherence preserved over the interrogation time $\tau = 2.5$ sec; and (c) loss of coherence with the Zeeman splitting tuned nearly on resonance with the ion micromotion of 500 kHz.

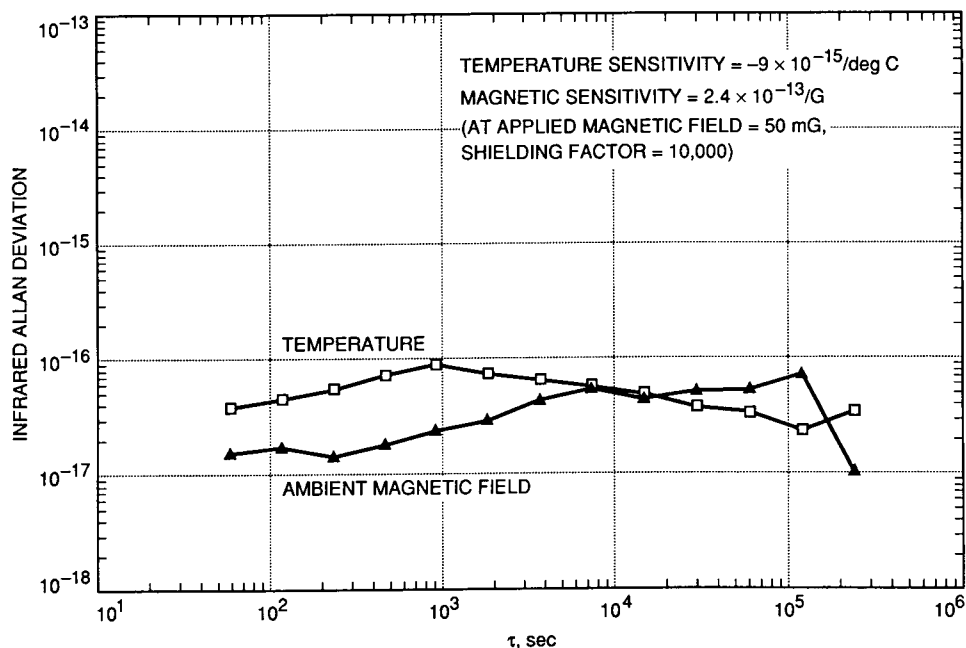


Fig. 7. Inferred Allan deviation based on: (a) magnetic and (b) temperature fluctuations measured in one of the environmental control chambers at the JPL's FSL. The plots are scaled by using the measured environmental sensitivity of the trapped ion frequency standard.

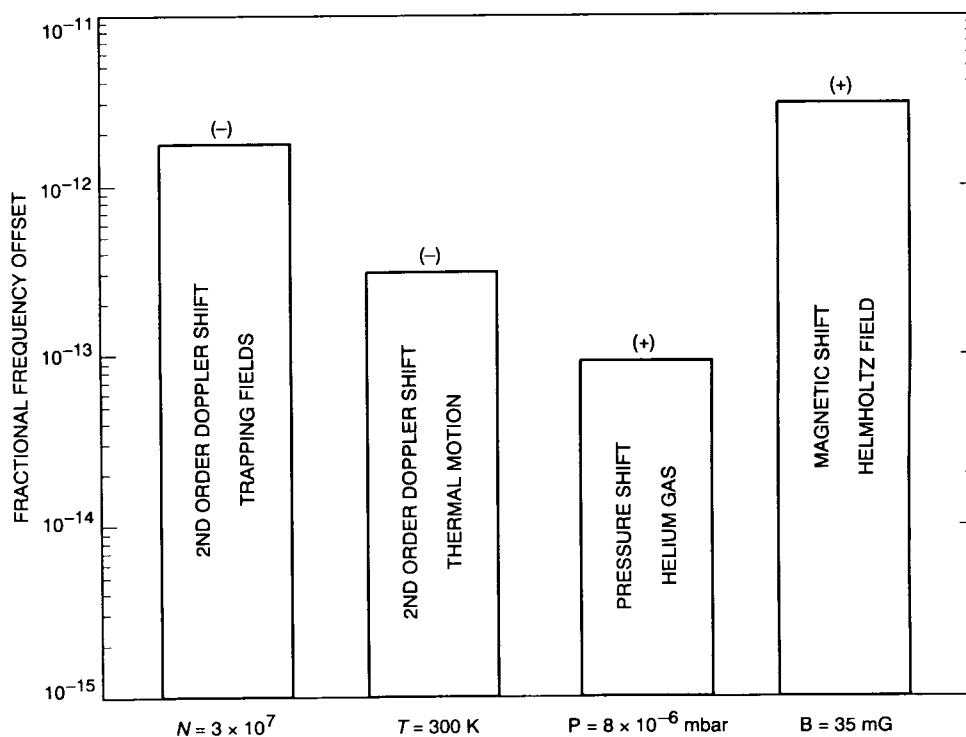


Fig. 8. The largest fractional frequency offsets to the $^{199}\text{Hg}^+$ clock transition in the linear trap under present operating conditions.

N92-24305

P-18

J574050

Orbit-Determination Performance of Doppler Data for Interplanetary Cruise Trajectories Part I: Error Analysis Methodology

J. S. Ulvestad and S. W. Thurman
Navigation Systems Section

This article describes an error covariance analysis methodology used to investigate different weighting schemes for two-way (coherent) Doppler data in the presence of transmission-media and observing-platform calibration errors. The analysis focuses on orbit-determination performance in the interplanetary cruise phase of deep-space missions. Analytical models for the Doppler observable and for transmission-media and observing-platform calibration errors are presented, drawn primarily from previous work. Previously published analytical models have been improved upon by (1) considering the effects of errors in the calibration of radio signal propagation through the troposphere and ionosphere as well as station-location errors; (2) modelling the spacecraft state transition matrix using a more accurate piecewise-linear approximation to represent the evolution of the spacecraft trajectory; and (3) incorporating Doppler data weighting functions that are functions of elevation angle, which reduce the sensitivity of the estimated spacecraft trajectory to troposphere and ionosphere calibration errors. The analysis is motivated by the need to develop suitable weighting functions for two-way Doppler data acquired at 8.4 GHz (X-band) and 32 GHz (Ka-band). This weighting is likely to be different from that in the weighting functions currently in use; the current functions were constructed originally for use with 2.3-GHz (S-band) Doppler data, which are affected much more strongly by the ionosphere than are the higher frequency data.

I. Introduction

Interplanetary spacecraft navigation is accomplished using a number of different techniques. The classical means of navigation have included those based on measurements of Doppler shift and on ranging measurements. In

recent years, differential very long baseline interferometry (VLBI) has been added to the arsenal of tools used for navigation [1]. For planetary approach and encounter, optical navigation using images of solar-system bodies against the background stars also is a very useful means of navigation [2]. Differenced (two-way minus three-way) Doppler

and range currently are being investigated for spacecraft tracking [3,4], as are other interferometric data types such as connected-element interferometry [5] and same-beam VLBI [6].

Despite the plethora of data types being developed, the workhorse navigation data type remains two-way (coherent) Doppler. These data are easy to acquire and can be obtained without interference with telemetry transmission from a spacecraft. However, fairly long tracking sessions ("passes") are necessary on each day of tracking. Doppler data are directly sensitive to motion along the line of sight; accumulation of Doppler data over a number of days provides increasing sensitivity to position and velocity in the plane of the sky (the plane perpendicular to the Earth-spacecraft line-of-sight). VLBI data yielding similar or improved *angular* accuracies often can be acquired in far less time at a greater operational cost and at the expense of interfering with telemetry transmission from the spacecraft. As instrumentation noise has been reduced and the calibration of various effects that cause systematic measurement errors has advanced, the orbit-determination performance obtained from Doppler data has improved substantially.

Obtaining the best accuracy from Doppler data often has required that an empirical trade-off be made between acquiring the most data possible in a given pass and down-weighting or deleting the data which can contain large errors due to imperfect calibration. In particular, it has been a common practice to deweight or discard data acquired at low elevation angles because of errors in the calibration of propagation effects in the Earth's ionosphere and troposphere. For Doppler data taken near 2.3 GHz (S-band), ionosphere calibration errors typically dominate the low-elevation data. Current missions often use the higher radio frequency of 8.4 GHz (X-band), and future missions may use transmission frequencies near 32 GHz (Ka-band) for improved telemetry performance. At these higher frequencies, the ionosphere calibration errors become less important (ionospheric propagation errors scale with the inverse square of the radio frequency), and troposphere calibration uncertainties dominate the errors in the low-elevation data. For Doppler data used in orbit determination, the weighting function currently implemented at JPL was established when radio metric data were obtained at 2.3 GHz and emphasized the minimization of errors caused by imperfect ionosphere calibration. Therefore this function is inappropriate for 8.4-GHz and 32-GHz data.

This article describes the development of a numerical error covariance analysis used to investigate the benefits of new elevation-dependent weighting schemes for Doppler

data acquired at high radio frequencies. The analysis described below is applicable to interplanetary cruise trajectories, but does not include features to enable computation of gravitational effects encountered during planetary approach or encounter phases. A piecewise-linear approximation to the evolution of the spacecraft state vector enables the analysis to be applied to data arcs that are months in length; keeping quadratic terms could enable the study of even longer data arcs at the expense of much additional computation. The virtue of the piecewise-linear analysis is that it is relatively simple, enabling consideration of a large number of different cases in a short period of time.

Previous analytical and numerical studies of Doppler-based orbit-determination performance can be found in a number of different references. For example, Hamilton and Melbourne [7] and Thurman [8] have considered the information contained in a single pass of Doppler measurements. Analyses have been extended to several days of data in a number of reports (e.g., [9-11]). However, most of these analyses have considered only data noise and, in some cases, station-location errors. The analysis and computational procedures described here are based on that previous work, but incorporate a number of improvements. A better approximation has been used for the propagation of observables from observation times to the reference epoch. Some terms that have been neglected in the past have been included. Effects of the errors in calibration of the static troposphere, the ionosphere, and the station locations all have been considered simultaneously, with special attention paid to the accuracy of the troposphere and ionosphere models. Finally, the methodology includes the ability to use different Doppler data-weighting functions in order to enable investigation of the improvement in orbit-determination accuracy that might be achieved by using different weighting schemes.

II. Summary of Least-Squares Error Covariance Analysis

The description of least-squares covariance analysis is kept to a minimum here; the reader should consult other references (e.g., [7,8,12]) for more explanation. Consider the range-rate $\dot{\rho}_i$ associated with the i th Doppler measurement. This measurement is modeled as a function of the epoch trajectory \vec{r}_0 , the measurement time t_i , and the random measurement error n_i , as follows:

$$\dot{\rho}_i = f(\vec{r}_0, t_i) + n_i \quad (1)$$

In a linearized model, measurement residuals $\Delta\dot{\rho}_i$ can be related to perturbations $\Delta\vec{r}_0$ in the trajectory by

$$\Delta \dot{\rho}_i \approx \frac{\partial \dot{\rho}_i}{\partial \vec{r}_0} \Delta \vec{r}_0 + n_i \quad (2)$$

The accumulated information matrix, J , for a set of N independent measurements of $\dot{\rho}_i$ with individual weights w_i is

$$J = \sum_{i=1}^N w_i (\partial \dot{\rho}_i / \partial \vec{r}_0)^T (\partial \dot{\rho}_i / \partial \vec{r}_0) \quad (3)$$

If each data point were weighted equally according to an estimate of the variance σ_D^2 of the Doppler measurement, w_i would be replaced by σ_D^{-2} , which could be taken outside the summation. However, since a primary goal of this investigation is to study the effect of a variety of data-weighting schemes, the more general formulation of Eq. (3) will be maintained.

Equation (3) holds for points within a continuous tracking pass as well as for points acquired over several tracking passes. If it is assumed that no a priori information is available, the error covariance matrix Λ_2 is the inverse of the information matrix:

$$\Lambda_2 = J^{-1} \quad (4)$$

III. Partial Derivatives and the Computed Doppler Error Covariance

Consider the case of a spacecraft whose right ascension and declination are denoted by (α, δ) , viewed from a tracking station whose cylindrical coordinates with respect to the Earth's center are (r_s, λ_s, z_s) . The station-spacecraft tracking geometry is shown in Fig. 1. The range-rate for the spacecraft, $\dot{\rho}$, which can be related directly to the measured Doppler shift, is given approximately by [10]

$$\begin{aligned} \dot{\rho} \approx & \dot{r} - z_s \dot{\delta} \cos \delta + r_s (\dot{\varphi} - \dot{\alpha}) \cos \delta \sin (\varphi - \alpha) \\ & + r_s \dot{\delta} \sin \delta \cos (\varphi - \alpha) \end{aligned} \quad (5)$$

In Eq. (5), r is the spacecraft distance from the center of the Earth, φ (Local Sidereal Time) is the angle between the station meridian and the vernal equinox, and time derivatives are indicated by dots above the variables. The quantity $(\varphi - \alpha)$ is the hour angle of the spacecraft, measured west from the local meridian. The partial derivatives of the range-rate with respect to the six-parameter state vector $\vec{r} = (r, \delta, \alpha, \dot{r}, \dot{\delta}, \dot{\alpha})$ are given by

$$A_1 \equiv \begin{pmatrix} \partial \dot{\rho} / \partial r \\ \partial \dot{\rho} / \partial \delta \\ \partial \dot{\rho} / \partial \alpha \\ \partial \dot{\rho} / \partial \dot{r} \\ \partial \dot{\rho} / \partial \dot{\delta} \\ \partial \dot{\rho} / \partial \dot{\alpha} \end{pmatrix} = \begin{pmatrix} 0 \\ z_s \dot{\delta} \sin \delta - r_s (\dot{\varphi} - \dot{\alpha}) \sin \delta \sin (\varphi - \alpha) + r_s \dot{\delta} \cos \delta \cos (\varphi - \alpha) \\ -r_s (\dot{\varphi} - \dot{\alpha}) \cos \delta \cos (\varphi - \alpha) + r_s \dot{\delta} \sin \delta \sin (\varphi - \alpha) \\ 1 \\ -z_s \cos \delta + r_s \sin \delta \cos (\varphi - \alpha) \\ -r_s \cos \delta \sin (\varphi - \alpha) \end{pmatrix} \quad (6)$$

To relate changes in the measured Doppler frequency to changes in the spacecraft state vector \vec{r}_0 at some reference epoch t_0 , one must find the spacecraft state transition matrix, consisting of the partial derivatives of the state vector \vec{r} with respect to the epoch state vector $\vec{r}_0 = (r_0, \delta_0, \alpha_0, \dot{r}_0, \dot{\delta}_0, \dot{\alpha}_0)$. These partial derivatives constitute a 6×6 matrix A_2 that is a function of the geometry at the reference time and the interval that has passed since that reference time. Formally, the matrix A_2 is expressed as $\partial \vec{r} / \partial \vec{r}_0$. The most common method of deriving this matrix analytically (e.g., [9]) is to linearize all six components of the state vector in time t as follows:

$$\left. \begin{aligned} r &= r_0 + \dot{r}_0 t \\ \delta &= \delta_0 + \dot{\delta}_0 t \\ \alpha &= \alpha_0 + \dot{\alpha}_0 t \\ \dot{r} &= \dot{r}_0 + \ddot{r}_0 t \\ \dot{\delta} &= \dot{\delta}_0 + \ddot{\delta}_0 t \\ \dot{\alpha} &= \dot{\alpha}_0 + \ddot{\alpha}_0 t \end{aligned} \right\} \quad (7)$$

The acceleration terms \ddot{r}_0 , $\ddot{\delta}_0$, and $\ddot{\alpha}_0$ in Eq. (7) can be expressed as functions of \ddot{r}_0 , thereby enabling a determination of all the relevant partial derivatives. The effects of gravitational acceleration due to the Sun are included in the computation of the partials. The gravitational effects of other bodies have been neglected, so the analysis is useful only for the interplanetary cruise portion of a trajectory, not for the planetary approach/encounter phase.

Formulas for the entries in the matrix A_2 have been derived in [8] and [10]. Occasional sign errors occur in the literature; correct expressions for these partials are given in the Appendix. The formulation here remains cast in terms of the state vector of the spacecraft instead of the coefficients (a,b,c,d,e,f) defined in [6] and [10] to be functions of the six parameters in the state vector.

Most previous covariance analyses (e.g., [9]) employed an A_2 matrix whose terms were functions only of the state vector \vec{r}_0 at the epoch time and the time interval since that epoch time. Such a formulation is a good approximation for multiple passes to the extent that the spacecraft motion fits the linear model over the entire data arc.

A more accurate representation of the partial derivatives indicated in Eq. (3) can be obtained by using a piecewise-linear approximation. For each pass, an epoch time is chosen to be the time of spacecraft transit *during that pass*. A linear formulation similar to Eqs. (7) then is used to relate the state vector at any time during a pass to an epoch state vector for the same pass, using the parameter values for the epoch time of the given pass, *not* those for the epoch time of the first pass. Multiplication by the mapping matrices relating each day to the previous day, with the appropriate epoch state vector used for each one-day step in the propagation, provides the mapping matrix to the reference epoch. Hence, the continuously curving spacecraft trajectory is approximated by a series of (six-dimensional) straight lines of different slope rather than by a single straight line.

Mathematically, let the matrix of partials of the state vector for pass j at time t with respect to the epoch state vector at the midpoint of pass j be denoted by $M_j(t)$. This matrix has the same entries as the matrix A_2 (results given in the Appendix), except that \vec{r}_0 is replaced by $\vec{r}_{j,0}$, the spacecraft state vector at the transit time for pass j , and the time t is referred to the reference time for pass j , not to the first pass. To relate measurements at time t during pass j to the epoch state vector on day 1, the matrix A_2 should be replaced as follows:

$$\begin{aligned} A_2 \rightarrow B_2(t) &= M_j(t) A_j^{j-1} A_{j-1}^{j-2} \cdots A_3^2 A_2^1 \\ &= M_j(t) \prod_{m=2}^j A_{j+2-m}^{j+1-m} \end{aligned} \quad (8)$$

The matrix A_j^{j-1} represents the mapping matrix from the midpoint of pass j to the midpoint of pass $j-1$. Computation of A_j^{j-1} is similar to the computation of the matrix A_2 . The straightforward method is to use a linearization process similar to that of Eqs. (7):

$$\left. \begin{aligned} r_{j,0} &= r_{j-1,0} + \tau \dot{r}_{j-1,0} \\ \delta_{j,0} &= \delta_{j-1,0} + \tau \dot{\delta}_{j-1,0} \\ \alpha_{j,0} &= \alpha_{j-1,0} + \tau \dot{\alpha}_{j-1,0} \\ \dot{r}_{j,0} &= \dot{r}_{j-1,0} + \tau \ddot{r}_{j-1,0} \\ \dot{\delta}_{j,0} &= \dot{\delta}_{j-1,0} + \tau \ddot{\delta}_{j-1,0} \\ \dot{\alpha}_{j,0} &= \dot{\alpha}_{j-1,0} + \tau \ddot{\alpha}_{j-1,0} \end{aligned} \right\} \quad (9)$$

In Eqs. (9), $\vec{r}_{j,0}$ is the state vector of the spacecraft at the transit time for pass j , $\vec{r}_{j-1,0}$ is the similar state vector for the previous pass, and τ is the time between the two transits. In the simple piecewise-linear approximation, A_j^{j-1} would have entries similar to those for A_2 (see Appendix), except that \vec{r}_0 would be replaced by $\vec{r}_{j-1,0}$ and t would be replaced by τ . A slightly better approximation has been used in the analysis described here: The entries of A_j^{j-1} are evaluated at $\vec{r}_{j,j-1} \equiv 0.5(\vec{r}_{j-1,0} + \vec{r}_{j,0})$ instead of at $\vec{r}_{j-1,0}$. Thus, the propagation from the transit time for pass j to the transit time for pass $j-1$ uses the average slope of the six-dimensional state vector over the time period between these two times rather than the slope at either end of the period.

Figure 2 is a schematic of several methods of propagating data to the reference time; the curvature of a trajectory over 10 days (i.e., the acceleration) has been exaggerated to highlight the differences. If the actual evolution of a spacecraft state vector is represented by the lower quadratic curve D, the propagation of a state vector found for day 10 to the reference time on day 1 can be represented by three different approximations to the slope of that curve. The horizontal line (labelled A) is the result of using the slope on day 1 to estimate the state vector that would be found for day 1 by propagating back in time from the state vector actually observed on day 10; this is

roughly equivalent to propagating observations from day 10 to the reference time on day 1, using only the epoch state vector of day 1 to evaluate the matrix A_2 . The uppermost piecewise-linear "trajectory" (labelled B) uses the slope on day $j-1$ to propagate the trajectory from day j to day $j-1$; it is the same (in principle) as evaluating each matrix A_j^{j-1} at the point $\vec{r}_{j-1,0}$ (namely, the spacecraft state vector at the midpoint of pass $j-1$) and performing the calculation indicated by Eq. (8). Another piecewise-linear approximation (labelled C), virtually indistinguishable from the actual quadratic trajectory (labelled D) in Fig. 2, is equivalent to evaluating each A_j^{j-1} at $\vec{r}_{j,j-1}$ (defined above) and performing the multiplications in Eq. (8). Clearly, the piecewise-linear approximation described in connection with Eqs. (8) and (9) is the best of the three methods shown, although the degree to which it surpasses the other methods depends on the actual curvature of the spacecraft trajectory over a particular data arc.

The partial derivatives of each Doppler measurement with respect to the epoch state vector now can be determined readily from

$$\frac{\partial \dot{\rho}_i}{\partial \vec{r}_0} = \frac{\partial \dot{\rho}_i}{\partial \vec{r}} B_2(t) \quad (10)$$

where B_2 is evaluated as shown in Eq. (8) and described in the accompanying text. Using Eqs. (3), (4), and (10), the error covariance matrix associated with a weighted least-squares estimate of \vec{r}_0 can be calculated.

It often is convenient to express the epoch state error covariance in a Cartesian coordinate frame rather than with spherical coordinates. A particularly useful Cartesian frame is the plane-of-sky coordinate system defined by the spacecraft position at epoch, which is depicted in Fig. 3. The position components r_c , r_δ , and r_α represent small position perturbations in the direction of increasing r , δ , and α coordinate directions relative to the nominal epoch spacecraft position. To make the conversion, it is necessary to compute the matrix A_3 containing the partial derivatives of the Cartesian coordinates at the reference time, $(r_{c,0}, r_{\delta,0}, r_{\alpha,0}, \dot{r}_{c,0}, \dot{r}_{\delta,0}, \dot{r}_{\alpha,0})$, with respect to the spherical coordinates $(r_0, \delta_0, \alpha_0, \dot{r}_0, \dot{\delta}_0, \dot{\alpha}_0)$. Reference [11] gives those derivatives, but contains several errors; correct values for the partials matrix A_3 are given in the Appendix.

It easily can be shown that the plane-of-sky Cartesian covariance matrix, Λ_{pos} , is given by

$$\Lambda_{\text{pos}} = A_3 \Lambda_{\text{sph}} A_3^T \quad (11)$$

In Eq. (11), Λ_{sph} is simply the error covariance matrix for the spacecraft spherical coordinates at epoch.

IV. Description of Consider-State Error Covariance Analysis

A. Brief Review of Theory

The error covariance calculation outlined in Sections II and III could be used to predict the orbit-determination error statistics obtained from a series of Doppler measurements if each measurement were affected only by random, uncorrelated errors. However, systematic errors also can be present in the Doppler data due to imperfect calibration of a number of quantities that often are not solved for in the orbit-determination process. The sensitivity of the computed trajectory to these error sources can be estimated by allowing their influence to be "considered" in determining the error statistics associated with the computed trajectory. Obviously, the use of such a procedure results in a suboptimal estimate of the spacecraft trajectory, an estimate that does not minimize the expected value of the mean-squared estimation errors. This subsection reviews briefly the procedures for calculating a "consider" error covariance, while the following subsection describes the particular set of considered parameters included in the current analysis.

A consider-state covariance analysis primarily involves the computation of the partial derivatives of the estimated epoch state vector, $\vec{r}_{0,e}$, with respect to the consider parameters \vec{y}_c . The consider parameters are those quantities that are not estimated explicitly in the orbit determination. To perform the computation, the partial derivatives of the Doppler observable $\dot{\rho}$ with respect to the consider parameters \vec{y}_c must be found and then used in combination with the partials of $\dot{\rho}$ with respect to \vec{r}_0 to calculate the desired quantities. (Note that $\vec{r}_{0,e}$ is the *estimated* state vector, while \vec{r}_0 is the *actual* state vector.) Formally, define the matrix $\partial \vec{r}_{0,e} / \partial \vec{y}_c$ as the sensitivity matrix, denoted as S . A detailed derivation of the sensitivity matrix is given by McReynolds [13]; only the results will be summarized here. The sensitivity matrix can be expressed as a function of the computed covariance, Λ_2 , representing the uncertainty in the estimated parameters due to the random measurement errors (the "data noise"), and an additional matrix as follows:

$$\left. \begin{aligned} S &\equiv \partial \vec{r}_{0,e} / \partial \vec{y}_c \\ &= \Lambda_2 \sum_{i=1}^N w_i (\partial \dot{\rho}_i / \partial \vec{r}_0)^T (\partial \dot{\rho}_i / \partial \vec{y}_c) \end{aligned} \right\} \quad (12)$$

Recall that the computed covariance, Λ_2 , given in Eq. (4) is just the inverse of the information matrix. Now let the matrix Λ_y represent the error covariance matrix for the consider parameters. The matrix Λ_y is assumed to be diagonal in this analysis, with entries equal to the variances in the knowledge of the individual consider parameters. The statistical uncertainty introduced into the estimated parameters (the spacecraft epoch state vector elements) by the consider parameters, denoted by the covariance matrix Λ_1 , is given by

$$\Lambda_1 = S\Lambda_y S^T \quad (13)$$

Finally, the total, or consider, error covariance for the estimated parameters, Λ_3 , due to both random measurement errors and the consider parameters, is

$$\Lambda_3 = \Lambda_1 + \Lambda_2 \quad (14)$$

If desired, the error covariance matrix for the spacecraft epoch state vector given in Eq. (14) can be expressed readily in a plane-of-sky Cartesian frame using the mapping formula in Eq. (11).

The formulation given thus far for the consider error covariance is valid only if the weights w_i assigned to the Doppler data are equal to $1/\sigma_D^2$. If the weights used are not the inverse of the data noise variance, then the formulas given for Λ_1 and Λ_2 in Eqs. (13) and (4) must be modified. This scenario is of particular interest here, since the following section describes a Doppler weighting technique in which the data weights are *not* equal to the inverse of the data noise variance. This weighting scheme is designed to reduce the sensitivity of Doppler-based trajectory estimates to errors in the calibration of the transmission media and the station locations. For this more general problem, it can be shown¹ that the computed covariance, Λ_2 , is

$$\Lambda_2 = (A^T W A)^{-1} A^T W \Lambda_y W A (A^T W A)^{-1} \quad (15a)$$

where

$$A = \begin{pmatrix} \partial \dot{\rho}_1 / \partial \vec{r}_0 \\ \partial \dot{\rho}_2 / \partial \vec{r}_0 \\ \vdots \\ \partial \dot{\rho}_N / \partial \vec{r}_0 \end{pmatrix} \quad (15b)$$

$$W = \begin{pmatrix} w_1 & 0 & \cdot & \cdot & 0 \\ 0 & w_2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & 0 \\ 0 & 0 & \cdot & \cdot & w_N \end{pmatrix} \quad (15c)$$

and

$$\Lambda_y = I \cdot \sigma_D^2 \quad (15d)$$

For $W = I \cdot \sigma_D^{-2}$, Eq. (15a) reduces to $\Lambda_2 = (A^T W A)^{-1}$, as it should. To form the sensitivity matrix in the general case, Eq. (12) still is appropriate, but with the expression in Eq. (15a) used for the computed covariance Λ_2 . The formula for the consider error covariance Λ_3 , given in Eq. (14), then is unchanged after making the necessary modification to Eq. (12). This general formulation was used in the analysis described herein.

B. Consider Partial Derivatives

The particular analysis described here includes three different consider error sources. These consider parameters are (1) the calibration errors in the wet and dry components of the static zenith tropospheric delay; (2) the calibration errors in the daytime and nighttime ionospheric delay; and (3) the uncertainties in the station location. All delays are expressed in units of length, and fluctuations in the delays are ignored here. Although the uncertainties in universal time and polar motion (UTPM) are not included explicitly, these errors give signatures similar to uncertainties in station location; for instance, an error in universal time is exactly equivalent to an error in station longitude. Therefore, the analysis should give a reasonable picture of the effects of UTPM errors provided that the assumed magnitudes of the station-location errors are large enough to encompass the UTPM errors as well.

The partial derivatives of the Doppler observable with respect to calibration errors in the wet and dry components of the zenith troposphere have been computed by Russell² and are reproduced here. Define γ to be the elevation angle, and $\dot{\gamma}$ to be its time derivative. Then, for a single tracking station, the partial derivative with respect to an error $\Delta \rho_{zw}$ in the calibration of the wet zenith delay is given by

$$\frac{\partial \dot{\rho}}{\partial (\Delta \rho_{zw})} = 2f(\mathcal{A}_2, \mathcal{B}_2, \gamma, \dot{\gamma}) \quad (16a)$$

¹ R. K. Russell, "The Development of Parameter Estimation," Engineering Note 001-100 (internal document), Jet Propulsion Laboratory, Pasadena, California, January 19, 1987.

² R. K. Russell, "Computation of Troposphere Partial Derivatives," JPL Engineering Memorandum 391-277 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 3, 1972.

where

$$f(\mathcal{A}, \mathcal{B}, \gamma, \dot{\gamma}) = -\dot{\gamma} \left[\sin \gamma + \frac{\mathcal{A}}{\tan \gamma + \mathcal{B}} \right]^{-2} \times \left[\cos \gamma - \frac{\mathcal{A}}{(\sin \gamma + \mathcal{B} \cos \gamma)^2} \right] \quad (16b)$$

The partial with respect to an error $\Delta\rho_{zD}$ in the calibration of the dry zenith delay is

$$\frac{\partial \dot{\rho}}{\partial (\Delta\rho_{zD})} = 2f(\mathcal{A}_1, \mathcal{B}_1, \gamma, \dot{\gamma}) \quad (17)$$

where f was defined in Eq. (16b). The constants \mathcal{A}_1 , \mathcal{A}_2 , \mathcal{B}_1 , and \mathcal{B}_2 were determined by ray-tracing calculations to be

$$\left. \begin{aligned} \mathcal{A}_1 &= 0.00143 \\ \mathcal{A}_2 &= 0.00035 \\ \mathcal{B}_1 &= 0.0445 \\ \mathcal{B}_2 &= 0.017 \end{aligned} \right\} \quad (18)$$

The consider partials that relate to errors in calibration of the ionospheric propagation have been given by Jacobson³ in terms of the daytime and nighttime "ionospheric coefficients" (\mathcal{D} and \mathcal{N}), which essentially represent the delay errors at zenith during the day and night. The expressions for the partial derivatives are

$$\frac{\partial \dot{\rho}}{\partial \mathcal{D}} = -\cos x \frac{d}{dt}(\sec z) - \dot{x} \sin x \sec z \quad (19a)$$

and

$$\frac{\partial \dot{\rho}}{\partial \mathcal{N}} = -\frac{d}{dt}(\sec z) \quad (19b)$$

In Eqs. (19a) and (19b), z is the zenith angle of the spacecraft as seen from the point where the line of sight from the tracking station passes through the mean ionospheric

height of 350 km ($z = \sin^{-1}[0.948 \cos \gamma]$); x is the minimum of $\pi/2$ and $|2\pi(t-\theta)/P|$, where P is the period of the diurnal variation of the ionosphere (taken to be 32 hours), θ is the phase of that variation (taken to be 14 hours of time), and t is the local time at the subionospheric point. To compute the partials, this time can be written in terms of the station location, the universal time, and the direction to the spacecraft; the interested reader is referred to the work by Jacobson³ or to [14] for further details.

Making use of Eq. (5) and the fact that for longitudes increasing to the east, $\partial/\partial\lambda_s = \partial/\partial\varphi$ (i.e., moving the station to the east is the same as increasing the local sidereal time), the partial derivatives with respect to the cylindrical station coordinates (r_s, λ_s, z_s) easily are shown to be

$$\left. \begin{aligned} \frac{\partial \dot{\rho}}{\partial r_s} &= \left. \begin{aligned} &(\dot{\varphi} - \dot{\alpha}) \cos \delta \sin (\varphi - \alpha) \\ &+ \dot{\delta} \sin \delta \cos (\varphi - \alpha) \end{aligned} \right\} \\ \frac{\partial \dot{\rho}}{\partial \lambda_s} &= \left. \begin{aligned} &r_s [(\dot{\varphi} - \dot{\alpha}) \cos \delta \cos (\varphi - \alpha) \\ &- \dot{\delta} \sin \delta \sin (\varphi - \alpha)] \end{aligned} \right\} \\ \frac{\partial \dot{\rho}}{\partial z_s} &= -\dot{\delta} \cos \delta \end{aligned} \right\} \quad (20)$$

In Eq. (20), the partials with respect to r_s and z_s are in units of velocity divided by distance, while the partial with respect to λ_s is in units of velocity per radian. To convert the partial with respect to λ_s to the same units as the other partials, the partial derivative given in Eq. (20) should be divided by r_s .

V. Weighting Functions

The consider covariance analysis procedure described above carried the weighting factor w_i for each Doppler measurement; w_i nominally is chosen to be $1/\sigma_D^2$ in conventional least-squares estimation. However, Doppler data can be affected by large consider errors at low elevations because of the effects of the ionosphere and the troposphere on signal propagation, as well as the elevation-independent effects of uncertain calibration of station locations and Earth-orientation parameters. In many cases, the inclusion of low-elevation data actually has resulted in larger uncertainties in operational navigation than have been obtained by discarding those data. Therefore, one common approach to this problem has been to discard all data below a specified elevation angle. A second approach has been to deweight all data uniformly, which

³ R. A. Jacobson, "Troposphere and Ionosphere Delay Models in ATHENA," JPL Engineering Memorandum 314-290 (internal document), Jet Propulsion Laboratory, Pasadena, California, November 16, 1982.

has the effect of making the high-elevation data "as bad" as the low-elevation data and is not likely to give the best trajectory determination. A third possibility, considered below, is to deweight the low-elevation data according to some function appropriate to minimize troposphere and/or ionosphere effects on the estimated trajectory. (This deweighting might be accompanied by imposition of an elevation cutoff.) The single most important motivation for constructing the consider analysis methodology described in this article was to make a systematic, quantitative investigation of various weighting functions that would lead to recommendations regarding appropriate methods of weighting 8.4-GHz and 32-GHz Doppler data.

The consider analysis described here has been implemented in a computer program that includes a weighting function of the following form:

$$w_i = \left[\sigma_D^2 + \left(\frac{\sigma_e}{(\sin \gamma)^q} \right)^2 \right]^{-1}, \quad \gamma > \gamma_{\min} \\ = 0, \quad \gamma < \gamma_{\min} \quad \left. \vphantom{\left[\sigma_D^2 + \left(\frac{\sigma_e}{(\sin \gamma)^q} \right)^2 \right]^{-1}} \right\} \quad (21)$$

This weighting function was postulated because it resembles, in an approximate form, the behavior of the troposphere partial derivatives given by Eqs. (16) and (17). As before, γ is the elevation angle of the spacecraft as viewed from the tracking station. The elevation-dependent part of the weighting function has the form $(\sigma_e / \sin^q \gamma)^2$, where q is a selectable exponent and σ_e is a selectable weighting factor. The assumed Doppler data noise variance is σ_D^2 , and all data at elevations below γ_{\min} are discarded (this elevation cutoff also is user-selectable). Note that with this form of weighting, the data analyst has the option of adjusting the weighting function in several ways. Data can be deweighted uniformly by increasing the value of σ_D and choosing $\sigma_e = 0$. Data can be deweighted as a function of elevation angle by appropriate selection of σ_e and q . Finally, γ_{\min} can be varied to change the cutoff angle.

Figures 4(a) and 4(b) are plots of the partial derivatives given by Eqs. (16)–(19) as functions of the elevation angle γ . These plots are for a segment of the Mars Observer trajectory, whose state vector is given in Table 1. Tracking from the Goldstone Deep Space Network complex (latitude of about 35.2 deg) is assumed. Note that the troposphere partials are dominant at the low elevation angles, reaching values about 40 times greater than the maximum ionosphere partials in this particular case. For

observations at 8.4 GHz, the uncertainties in the coefficients \mathcal{D} and \mathcal{N} are comparable to the values of $\Delta\rho_{zD}$ and $\Delta\rho_{zw}$, implying that the troposphere calibration error will dominate the effects of the ionosphere calibration error in the Doppler data. At 2.3 GHz, \mathcal{D} and \mathcal{N} are larger by a factor of ~ 13 and can be much more important.

Figure 5 shows the Doppler data weights as a function of elevation angle for several different values of the variable parameters in Eq. (21); of course, the weights at the lower elevation points can be set to zero arbitrarily by making a choice of γ_{\min} . In this figure, an arbitrary value of unity is assigned to the Doppler data noise (i.e., $\sigma_D \equiv 1$). Values of 1, 2, and 3 are used for q , while σ_e is set equal to either σ_D or $0.5 \sigma_D$. For comparison, Fig. 6 is a plot of the *inverse* of the troposphere partial derivatives. If the troposphere were the *only* contributing calibration error and dominated the data noise in determining the Doppler error, the ideal weighting function would have a shape similar to the curves in Fig. 6. A comparison of Figs. 5 and 6 implies that the weighting functions with values of $q = 2$ or $q = 3$ may be more appropriate for Doppler data whose errors are dominated by errors in the troposphere calibration.

Note that the shapes of the weighting functions in Fig. 5 already lead to an interesting conclusion. When the weighting function is a strong function of elevation angle (i.e., $q \approx 3$), the choice of elevation cutoffs should make very little difference in the final results of the analysis. This is because, for the larger values of q , the low-elevation data already are downweighted so much that they have little effect on the orbit determination; they effectively have been discarded even before the elevation limit is applied. The specification of an elevation cutoff would be more important if the ratio of σ_e to σ_D in Eq. (21) were much smaller than the values used in generating the plot in Fig. 5.

VI. Accuracy Over Long Data Arcs

It was mentioned previously that the trajectory representation developed herein breaks down for long data arcs, but the maximum length of time for which the approximations are valid was not specified. However, it is possible to make a rough estimate of the longest data arcs that might be investigated using the formalism described above. This estimate has been made by considering the accuracy of the linear approximation given by Eqs. (7) and the accuracy of the piecewise-linear approximation described by Eqs. (8) and (9) and in the surrounding discussion.

A more accurate system of equations would include terms that are quadratic in time:

$$\left. \begin{aligned} r &= r_0 + \dot{r}_0 t + \ddot{r}_0 t^2/2 \\ \delta &= \delta_0 + \dot{\delta}_0 t + \ddot{\delta}_0 t^2/2 \\ \alpha &= \alpha_0 + \dot{\alpha}_0 t + \ddot{\alpha}_0 t^2/2 \\ \dot{r} &= \dot{r}_0 + \ddot{r}_0 t + t^2(\partial^3 r_0/\partial t^3)/2 \\ \dot{\delta} &= \dot{\delta}_0 + \ddot{\delta}_0 t + t^2(\partial^3 \delta_0/\partial t^3)/2 \\ \dot{\alpha} &= \dot{\alpha}_0 + \ddot{\alpha}_0 t + t^2(\partial^3 \alpha_0/\partial t^3)/2 \end{aligned} \right\} \quad (22)$$

As an example, consider the partial derivative $\partial r/\partial r_0$. In the linear theory, this partial derivative is found as follows:

$$\frac{\partial r}{\partial r_0} = 1 + t \frac{\partial \dot{r}_0}{\partial r_0} = 1 \quad (23a)$$

since $(\partial \dot{r}_0/\partial r_0) = 0$. Keeping the quadratic time terms as in Eq. (22), letting \ddot{r}_g be the line-of-sight acceleration due to the Sun's gravitational field, and using the result from [10] that

$$\ddot{r}_0 = r_0(\dot{\delta}_0^2 + \dot{\alpha}_0^2 \cos^2 \delta_0) + \ddot{r}_g \quad (23b)$$

this partial derivative is

$$\left. \begin{aligned} \frac{\partial r}{\partial r_0} &= 1 + t \frac{\partial \dot{r}_0}{\partial r_0} + \frac{t^2}{2} \frac{\partial \ddot{r}_0}{\partial r_0} \\ &= 1 + \frac{t^2}{2} \left(\dot{\delta}_0^2 + \dot{\alpha}_0^2 \cos^2 \delta_0 + \frac{\partial \ddot{r}_g}{\partial r_0} \right) \\ &= 1 + \frac{t^2}{2} \\ &\quad \times \left(\dot{\delta}_0^2 + \dot{\alpha}_0^2 \cos^2 \delta_0 + \frac{\mu}{r_{p,0}^3} [2 - 3 \sin^2 \psi_0] \right) \end{aligned} \right\} \quad (23c)$$

In the above equation, μ is the universal gravitational constant multiplied by the mass of the Sun, $r_{p,0}$ is the Sun-spacecraft distance, and ψ_0 is the Earth-spacecraft-Sun angle. The value of $\partial \ddot{r}_g/\partial r_0$ has been taken from [10]. As an example, a segment of the planned Mars Observer trajectory on July 22, 1993 was taken, with epoch state

parameters as shown in Table 1. Using the appropriate numbers, Eq. (23c) becomes

$$\frac{\partial r}{\partial r_0} = 1 + 2.6 \times 10^{-14} t^2 \quad (24)$$

where t is given in seconds. Thus, the linear approximation used in the standard derivation of the state transition matrix A_2 would have a fractional error of $2.6 \times 10^{-14} t^2$. The partial derivative would be accurate at the 1-percent level only if t were smaller than 6.2×10^5 seconds, or about 7 days. For 10-percent accuracy in the partial, t could be no greater than 23 days.

The piecewise-linear approximation to the trajectory includes the bulk of the effect of the quadratic terms. Considering the transition matrix connecting day 2 to day 1, $A_{2,1}^1$, the maximum value of t that would be used in Eq. (24) would be about a day, or 8.6×10^4 seconds. The connection from day 3 to day 1 is based on two values of the partial derivative, which differ because of the effects of the quadratic terms included in a piecewise-linear manner. Thus, the fractional error for each day's pass would be indicated by Eq. (24), but the total fractional error over an n -day data arc is approximately

$$\left. \begin{aligned} f_n &\approx (1 + 2.6 \times 10^{-14} T^2)^n - 1 \\ &\approx 2.6 \times 10^{-14} n T^2 \end{aligned} \right\} \quad (25a)$$

instead of

$$\left. \begin{aligned} f_n &\approx [1 + 2.6 \times 10^{-14} (nT)^2] - 1 \\ &= 2.6 \times 10^{-14} n^2 T^2 \end{aligned} \right\} \quad (25b)$$

where T is the length of a day in seconds. These two equations differ in only one respect, their dependence on the number of days n in the data arc. For $n = 7$, Eq. (25b) predicts a fractional error of 0.95 percent for the linear approximation, while Eq. (25a) predicts a fractional error of 0.14 percent for the piecewise-linear approximation. For a month-long data arc, $n = 30$, the fractional error for the linear approximation would be about 17 percent, while that for the piecewise-linear approximation would be only 0.6 percent! In fact, even this result is an overestimate of the inaccuracy of the piecewise-linear approximation used in the analysis developed in this article. When the transition state matrix between passes j and $j - 1$ is evaluated for the state vector $\vec{r}_{j,j-1}$ at the midpoint between the two

passes, a more appropriate value of T used in Eq. (25a) would be half the length of the day rather than the full length of a day. In any case, it is clear that the length of the data arc that can be considered is substantially longer than that possible for the simple linear approximation.

Provided that calculations of the neglected terms in other partial derivatives give results similar to those described above, for the particular sample trajectory considered here, the piecewise-linear approximation represented by Eq. (8) should be quite good for data arcs longer than a month. The approximation would hold for a longer period cruise trajectory in the outer solar system, say at the distance of Jupiter, since the numerical coefficient of T^2 in Eq. (25a) would be considerably smaller. For a trajectory near Venus, the gravitational acceleration due to the Sun would be larger, reducing the accuracy of the linear approximation. In that case, the advantage of the piecewise-linear approximation would be even more pronounced. It is possible to get arbitrarily close to the true model of state-vector evolution by breaking the data arc into more pieces, but the above estimate shows that the simple piecewise-linear approximation that is used should be quite good for most circumstances.

VII. Summary

This article has described an error covariance analysis methodology for Doppler measurements that can be used to predict errors in the estimated state vector of an interplanetary spacecraft in the cruise portion of its trajectory. The analysis and associated computer program

improve on previously published analyses in several ways: (1) The effects of imperfect calibration of the static troposphere, the static ionosphere, and the station location are all considered simultaneously; (2) the spacecraft state transition matrix is modelled more accurately, using a piecewise-linear approximation to represent the evolution of the spacecraft trajectory; and (3) the Doppler data can be weighted according to elevation angle in an attempt to minimize the covariance associated with determination of the spacecraft trajectory. The first and second "improvements" or better representations also are available in comprehensive software such as the JPL Orbit Determination Program [15], but the limited focus of the analysis method described in this article makes it preferable for investigation of a large number of possible cases of data weighting in a short period of time.

The analysis method described in this article is useful for estimating the errors in trajectory determination for data arcs longer than a month in the case of a spacecraft more than 1 AU from the Sun and far from the gravitational influence of any planets. In the case of spacecraft in the outer solar system (e.g., Voyager and Galileo), the reduced gravitational effect of the Sun and the smaller angular velocities of the spacecraft should enable use of this formulation for a considerably longer data arc; the period of validity needs to be evaluated on a case-by-case basis. It is likely that the accumulated effects of nongravitational, unmodelled accelerations would lead to greater errors than those caused by the piecewise-linear inclusion of the quadratic terms due to known gravitational effects.

Acknowledgments

The authors thank Jim McDanell for an enlightening discussion on modification of the consider error covariance for arbitrary data weighting, Stuart Demczak for supplying the state vector for the Mars Observer trajectory segment, and Bob Jacobson for careful review of a draft of this article.

References

- [1] J. K. Miller and K. H. Rourke, "The Application of Differential VLBI to Planetary Approach Orbit Determination," *Deep Space Network Progress Report*, vol. May-June 1977, Jet Propulsion Laboratory, Pasadena, California, pp. 84-90, August 15, 1977.
- [2] J. E. Riedel, W. M. Owen, Jr., J. A. Stuve, S. P. Synnott, and R. M. Vaughan, "Optical Navigation During the Voyager Neptune Encounter," paper AIAA-90-2877, AIAA/AAS Astrodynamics Conference, Portland, Oregon, August 20-22, 1990.
- [3] S. W. Thurman, "Deep-Space Navigation With Differenced Data Types, Part I: Differenced Range Information Content," *TDA Progress Report 42-103*, vol. July-September 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 47-60, November 15, 1990.
- [4] S. W. Thurman, "Deep-Space Navigation with Differenced Data Types, Part II: Differenced Doppler Information Content," *TDA Progress Report 42-103*, vol. July-September 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 61-69, November 15, 1990.
- [5] C. D. Edwards, "Angular Navigation on Short Baselines Using Phase Delay Interferometry," *IEEE Transactions on Instrumentation and Measurement*, vol. 38, no. 2, pp. 665-667, April 1989.
- [6] W. M. Folkner, D. B. Engelhardt, J. S. Border, and N. A. Mottinger, "Orbit Determination for Magellan and Pioneer 12 Using Same-Beam Interferometry," paper AAS 91-393, AIAA/AAS Astrodynamics Conference, Durango, Colorado, August 19-22, 1991.
- [7] T. W. Hamilton and W. G. Melbourne, "Information Content of a Single Pass of Doppler Data from a Distant Spacecraft," *JPL Space Programs Summary*, no. 37-39, vol. III, Jet Propulsion Laboratory, Pasadena, California, pp. 18-23, March-April 1966.
- [8] S. W. Thurman, "Comparison of Earth-Based Radio Metric Data Strategies for Deep Space Navigation," paper AIAA 90-2908, AIAA/AAS Astrodynamics Conference, Portland, Oregon, August 22, 1990.
- [9] D. W. Curkendall and S. R. McReynolds, "A Simplified Approach for Determining the Information Content of Radio Tracking Data," *Journal of Spacecraft and Rockets*, vol. 6, no. 5, pp. 520-525, May 1969.
- [10] V. J. Ondrasik and D. W. Curkendall, "A First-Order Theory for Use in Investigating the Information Content Contained in a Few Days of Radio Tracking Data," *Deep Space Network Progress Report*, JPL Technical Report 32-1526, vol. III, Jet Propulsion Laboratory, Pasadena, California, pp. 77-93, June 15, 1971.
- [11] D. W. Curkendall, "Problems in Estimation Theory with Applications to Orbit Determination," Ph.D. thesis, Los Angeles: University of California at Los Angeles, 1971.
- [12] P. R. Escobal, *Methods of Orbit Determination*, 2nd ed., Malabar, Florida: Robert E. Krieger Publishing Company, pp. 435-450, 1976.
- [13] S. R. McReynolds, "The Sensitivity Matrix Method for Orbit Determination Error Analysis, With Application to a Mars Orbiter," *JPL Space Programs Sum-*

mary, no. 37-56, vol. II, January-February 1969, Jet Propulsion Laboratory, Pasadena, California, pp. 85-87, March 31, 1969.

- [14] J. A. Klobuchar, *A First Order, World-Wide, Ionospheric Time-Delay Algorithm*, Document AFCRL-TR-75-0502, Hanscom AFB, Massachusetts: Air Force Cambridge Research Laboratories, September 25, 1975.
- [15] T. D. Moyer, *Mathematical Formulation of the Double-Precision Orbit Determination Program (DPODP)*, Technical Report 32-1527, Jet Propulsion Laboratory, Pasadena, California, May 15, 1971.

Table 1. Initial state vector for a segment of Mars Observer trajectory.

Parameter	Value
Year	1993
Day of year	203 (22 July)
Distance, km	3.1573×10^8
Right ascension (α), deg	169.0252
Declination (δ), deg	5.1308
Radial velocity, km/sec	11.5770
$d\alpha/dt$, deg/sec	6.2454×10^{-6}
$d\delta/dt$, deg/sec	-2.5866×10^{-6}

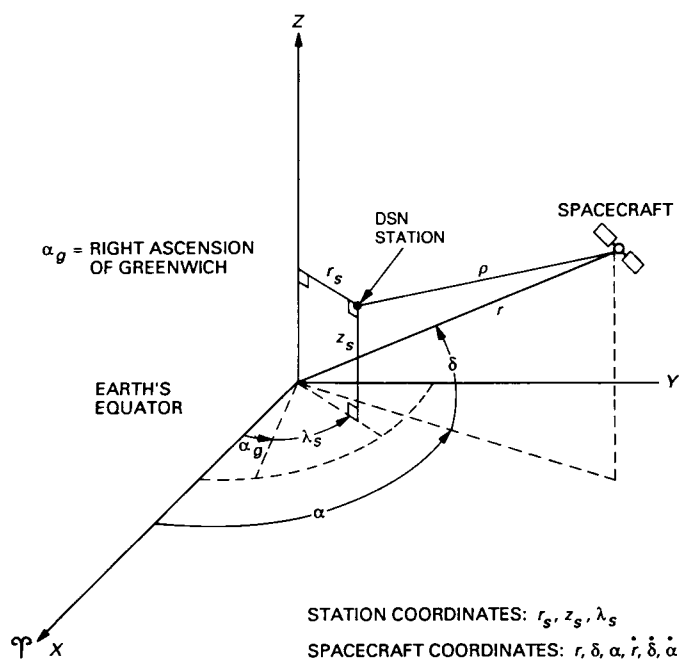


Fig. 1. Station-spacecraft tracking geometry.

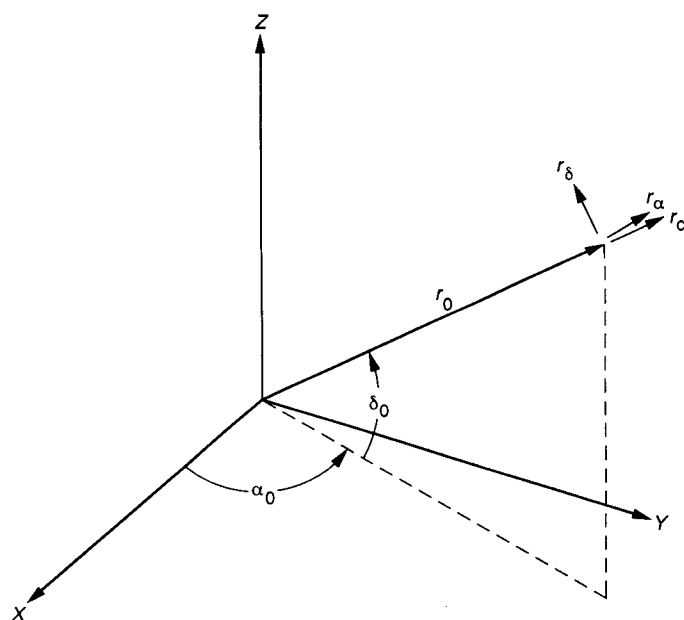


Fig. 3. Geocentric plane-of-sky Cartesian coordinate system.

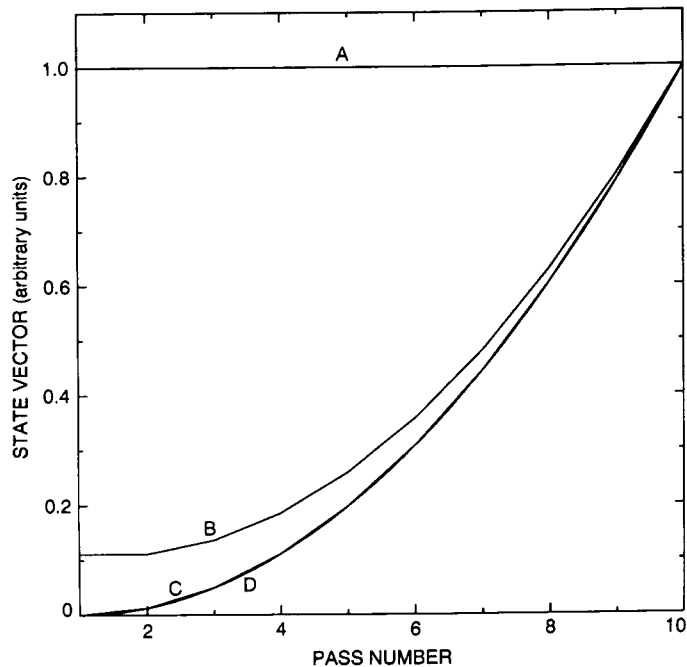


Fig. 2. The effect of different methods of computing a spacecraft state transition matrix.

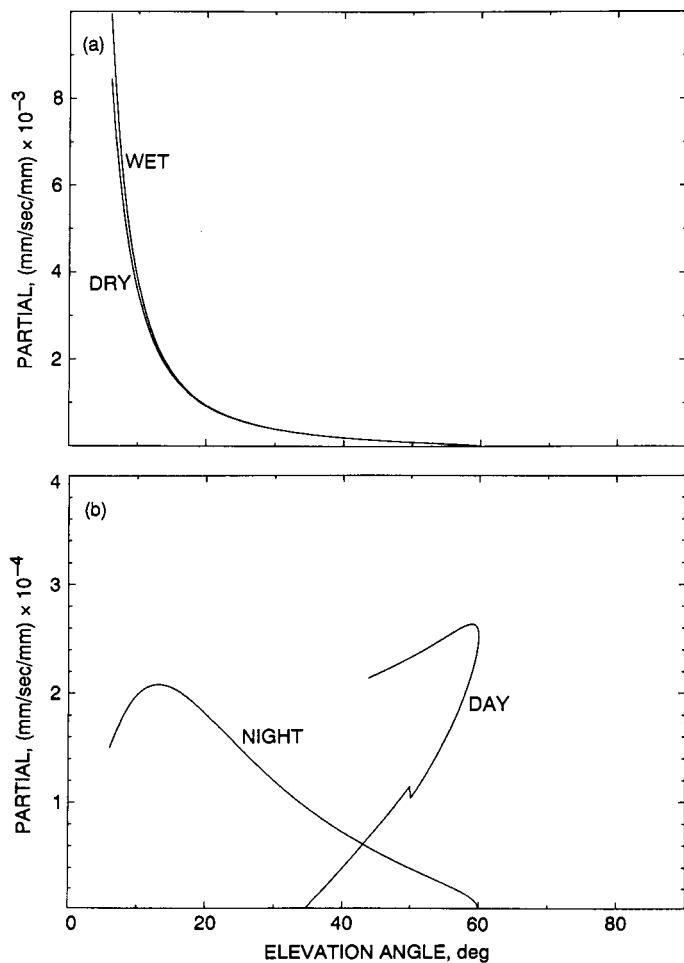


Fig. 4. Partial derivatives of the Doppler error with respect to the calibration errors in the troposphere and ionosphere delays plotted against elevation angle: (a) partials of the Doppler observable with respect to wet and dry zenith-troposphere calibration errors, and (b) partials of the Doppler observable with respect to daytime and nighttime ionosphere coefficients.

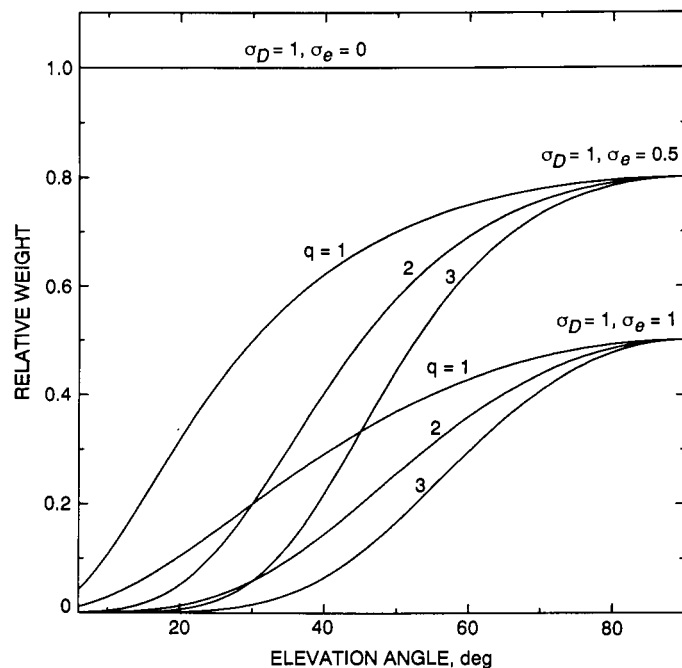


Fig. 5. The weighting function given by Eq. (21) for a cutoff elevation of 6 deg: Three different values of the exponent q are included for each family.

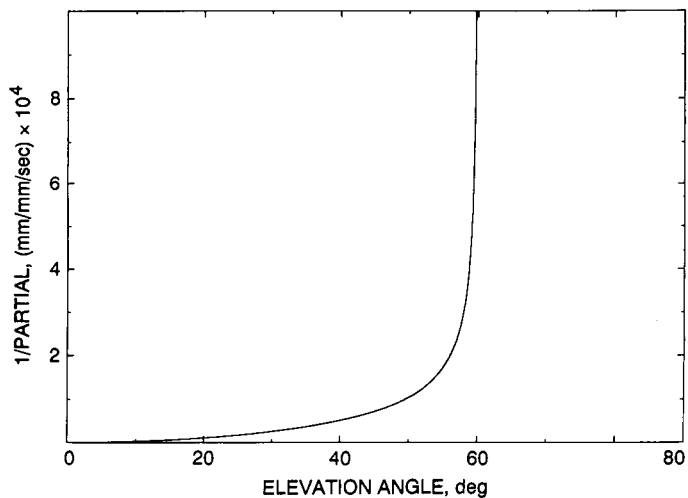


Fig. 6. Inverse of the partials of the Doppler error with respect to errors in the calibration of the zenith troposphere delay. Plots for wet and dry components overlap. The plot goes to infinity at transit, where γ goes to zero in Eq. (16b).

Appendix

Summary of Partial Derivatives

This appendix gives (without derivation) the partial derivatives in the 6×6 matrices A_2 and A_3 described in Section III. A_2 is the matrix of partials of the state vector \vec{r} at an arbitrary time t with respect to the state vector \vec{r}_0 at the epoch time for the same pass. To find the matrix of partials relating the state vector at the epoch time on day j to that at the epoch time on day $j-1$, denoted A_j^{j-1} in Section III, \vec{r}_0 should be replaced by $\vec{r}_{j,j-1}$ and t by τ , as described following Eq. (9).

The matrix A_2 is defined as follows:

$$A_2 = \begin{pmatrix} \frac{\partial r}{\partial r_0} & \frac{\partial r}{\partial \delta_0} & \frac{\partial r}{\partial \alpha_0} & \frac{\partial r}{\partial \dot{r}_0} & \frac{\partial r}{\partial \dot{\delta}_0} & \frac{\partial r}{\partial \dot{\alpha}_0} \\ \frac{\partial \delta}{\partial r_0} & \frac{\partial \delta}{\partial \delta_0} & \frac{\partial \delta}{\partial \alpha_0} & \frac{\partial \delta}{\partial \dot{r}_0} & \frac{\partial \delta}{\partial \dot{\delta}_0} & \frac{\partial \delta}{\partial \dot{\alpha}_0} \\ \frac{\partial \alpha}{\partial r_0} & \frac{\partial \alpha}{\partial \delta_0} & \frac{\partial \alpha}{\partial \alpha_0} & \frac{\partial \alpha}{\partial \dot{r}_0} & \frac{\partial \alpha}{\partial \dot{\delta}_0} & \frac{\partial \alpha}{\partial \dot{\alpha}_0} \\ \frac{\partial \dot{r}}{\partial r_0} & \frac{\partial \dot{r}}{\partial \delta_0} & \frac{\partial \dot{r}}{\partial \alpha_0} & \frac{\partial \dot{r}}{\partial \dot{r}_0} & \frac{\partial \dot{r}}{\partial \dot{\delta}_0} & \frac{\partial \dot{r}}{\partial \dot{\alpha}_0} \\ \frac{\partial \dot{\delta}}{\partial r_0} & \frac{\partial \dot{\delta}}{\partial \delta_0} & \frac{\partial \dot{\delta}}{\partial \alpha_0} & \frac{\partial \dot{\delta}}{\partial \dot{r}_0} & \frac{\partial \dot{\delta}}{\partial \dot{\delta}_0} & \frac{\partial \dot{\delta}}{\partial \dot{\alpha}_0} \\ \frac{\partial \dot{\alpha}}{\partial r_0} & \frac{\partial \dot{\alpha}}{\partial \delta_0} & \frac{\partial \dot{\alpha}}{\partial \alpha_0} & \frac{\partial \dot{\alpha}}{\partial \dot{r}_0} & \frac{\partial \dot{\alpha}}{\partial \dot{\delta}_0} & \frac{\partial \dot{\alpha}}{\partial \dot{\alpha}_0} \end{pmatrix} \quad (A-1)$$

A number of definitions must be made to write the partial derivatives in a reasonably compact form; the definitions made here largely follow those in [10], and all variables are assumed to be evaluated at the reference (or epoch) time indicated by the subscript "0." Let the right ascension and declination of the Sun be denoted by $\alpha_{s,0}$ and $\delta_{s,0}$, respectively. Denote the Earth-Sun distance by $r_{e,0}$ and the Earth-spacecraft distance by $r_{p,0}$. The Sun-Earth-spacecraft angle is χ_0 , where

$$\begin{aligned} \cos \chi_0 &= \cos \delta_0 \cos \delta_{s,0} \cos (\alpha_0 - \alpha_{s,0}) \\ &+ \sin \delta_0 \sin \delta_{s,0} \end{aligned} \quad (A-2)$$

The Earth-spacecraft-Sun angle ψ_0 is found from

$$\sin \psi_0 = r_{e,0} \sin \chi_0 / r_{p,0} \quad (A-3)$$

where

$$r_{p,0} = [r_0^2 + r_{e,0}^2 - 2r_0 r_{e,0} \cos \chi_0]^{1/2} \quad (A-4)$$

The angle σ_0 is given by

$$\begin{aligned} \cos \sigma_0 &= -\sin \delta_0 \cos \delta_{s,0} \cos (\alpha_0 - \alpha_{s,0}) \\ &+ \cos \delta_0 \sin \delta_{s,0} \end{aligned} \quad (A-5)$$

Reference [10] gives the gravitational accelerations of the spacecraft in the radial, declination, and right-ascension directions ($\ddot{r}_g, \ddot{\delta}_g, \ddot{\alpha}_g$). If μ is the product of the universal gravitational constant and the mass of the Sun, those accelerations are

$$\ddot{r}_g = \frac{-\mu r_0}{r_{p,0}^3} + \mu r_{e,0} \cos \chi_0 \left(\frac{1}{r_{p,0}^3} - \frac{1}{r_{e,0}^3} \right) \quad (A-6a)$$

$$\ddot{\delta}_g = \frac{\mu r_{e,0} \cos \sigma_0}{r_0} \left(\frac{1}{r_{p,0}^3} - \frac{1}{r_{e,0}^3} \right) \quad (A-6b)$$

$$\begin{aligned} \ddot{\alpha}_g &= \frac{-\mu r_{e,0} \cos \delta_{s,0} \sin (\alpha_0 - \alpha_{s,0})}{r_0} \\ &\times \left(\frac{1}{r_{p,0}^3} - \frac{1}{r_{e,0}^3} \right) \end{aligned} \quad (A-6c)$$

The nonzero partial derivatives in the matrix A_2 are given below:

$$\frac{\partial r}{\partial r_0} = 1 \quad (A-7a)$$

$$\frac{\partial r}{\partial \dot{r}_0} = t \quad (A-7b)$$

$$\frac{\partial \delta}{\partial \delta_0} = 1 \quad (A-8a)$$

$$\frac{\partial \delta}{\partial \dot{\delta}_0} = t \quad (A-8b)$$

$$\frac{\partial \alpha}{\partial \alpha_0} = 1 \quad (\text{A-9a})$$

$$\frac{\partial \alpha}{\partial \dot{\alpha}_0} = t \quad (\text{A-9b})$$

$$\frac{\partial \dot{r}}{\partial r_0} = \frac{\mu t (2 - 3 \sin^2 \psi_0)}{r_{p,0}^3} + (\dot{\delta}_0^2 + \dot{\alpha}_0^2 \cos^2 \delta_0) t \quad (\text{A-10a})$$

$$\begin{aligned} \frac{\partial \dot{r}}{\partial \delta_0} = & -2tr_0 \dot{\alpha}_0^2 \cos \delta_0 \sin \delta_0 + \mu tr_{e,0} \cos \sigma_0 \\ & \times \left(\frac{1}{r_{p,0}^3} - \frac{1}{r_{e,0}^3} - \frac{3r_0(r_0 - r_{e,0} \cos \chi_0)}{r_{p,0}^5} \right) \end{aligned} \quad (\text{A-10b})$$

$$\begin{aligned} \frac{\partial \dot{r}}{\partial \alpha_0} = & \mu tr_{e,0} \cos \delta_0 \cos \delta_{s,0} \sin (\alpha_0 - \alpha_{s,0}) \\ & \times \left(\frac{1}{r_{e,0}^3} - \frac{1}{r_{p,0}^3} + \frac{3r_0(r_0 - r_{e,0} \cos \chi_0)}{r_{p,0}^5} \right) \end{aligned} \quad (\text{A-10c})$$

$$\frac{\partial \dot{r}}{\partial \dot{r}_0} = 1 \quad (\text{A-10d})$$

$$\frac{\partial \dot{r}}{\partial \dot{\delta}_0} = 2r_0 \dot{\delta}_0 t \quad (\text{A-10e})$$

$$\frac{\partial \dot{r}}{\partial \dot{\alpha}_0} = 2r_0 \dot{\alpha}_0 t \cos^2 \delta_0 \quad (\text{A-10f})$$

$$\begin{aligned} \frac{\partial \dot{\delta}}{\partial r_0} = & \frac{2t\dot{r}_0 \dot{\delta}_0}{r_0^2} - \frac{t\ddot{\delta}_g}{r_0} \\ & + \frac{3\mu tr_{e,0} \cos \sigma_0}{r_{p,0}^5} \left(\frac{r_{e,0}}{r_0} \cos \chi_0 - 1 \right) \end{aligned} \quad (\text{A-11a})$$

$$\begin{aligned} \frac{\partial \dot{\delta}}{\partial \delta_0} = & -t\dot{\alpha}_0^2 (\cos^2 \delta_0 - \sin^2 \delta_0) - \frac{t\ddot{\delta}_g \cos \chi_0}{\cos \sigma_0} \\ & + \frac{3\mu tr_{e,0}^2 \cos^2 \sigma_0}{r_{p,0}^5} \end{aligned} \quad (\text{A-11b})$$

$$\begin{aligned} \frac{\partial \dot{\delta}}{\partial \alpha_0} = & \mu tr_{e,0} \cos \delta_{s,0} \sin (\alpha_0 - \alpha_{s,0}) \\ & \times \left[\frac{\sin \delta_0}{r_0} \left(\frac{1}{r_{e,0}^3} - \frac{1}{r_{p,0}^3} \right) - \frac{3r_{e,0}^2 \cos \delta_0 \cos \sigma_0}{r_{p,0}^5} \right] \end{aligned} \quad (\text{A-11c})$$

$$\frac{\partial \dot{\delta}}{\partial \dot{r}_0} = \frac{-2t\dot{\delta}_0}{r_0} \quad (\text{A-11d})$$

$$\frac{\partial \dot{\delta}}{\partial \dot{\delta}_0} = 1 - \frac{2t\dot{r}_0}{r_0} \quad (\text{A-11e})$$

$$\frac{\partial \dot{\delta}}{\partial \dot{\alpha}_0} = -2t\dot{\alpha}_0 \cos \delta_0 \sin \delta_0 \quad (\text{A-11f})$$

$$\begin{aligned} \frac{\partial \dot{\alpha}}{\partial r_0} = & \frac{2t\dot{r}_0 \dot{\alpha}_0}{r_0^2} - \frac{t\ddot{\alpha}_g}{r_0 \cos \delta_0} \\ & + \frac{3\mu tr_{e,0} \cos \delta_{s,0} \sin (\alpha_0 - \alpha_{s,0})(r_0 - r_{e,0} \cos \chi_0)}{r_0 r_{p,0}^5 \cos \delta_0} \end{aligned} \quad (\text{A-12a})$$

$$\begin{aligned} \frac{\partial \dot{\alpha}}{\partial \delta_0} = & \frac{t(2\dot{\delta}_0 \dot{\alpha}_0 + \ddot{\alpha}_g \sin \delta_0)}{\cos^2 \delta_0} \\ & - \frac{3\mu tr_{e,0}^2 \cos \sigma_0 \cos \delta_{s,0} \sin (\alpha_0 - \alpha_{s,0})}{r_{p,0}^5 \cos \delta_0} \end{aligned} \quad (\text{A-12b})$$

$$\begin{aligned} \frac{\partial \dot{\alpha}}{\partial \alpha_0} = & \frac{t\ddot{\alpha}_g}{\cos \delta_0 (\tan \alpha_0 - \tan \alpha_{s,0})} \\ & + \frac{3\mu tr_{e,0}^2 \cos^2 \delta_{s,0} \sin^2 (\alpha_0 - \alpha_{s,0})}{r_{p,0}^5} \end{aligned} \quad (\text{A-12c})$$

$$\frac{\partial \dot{\alpha}}{\partial \dot{r}_0} = \frac{-2t\dot{\alpha}_0}{r_0} \quad (\text{A-12d})$$

$$\frac{\partial \dot{\alpha}}{\partial \dot{\delta}_0} = 2t\dot{\alpha}_0 \tan \delta_0 \quad (\text{A-12e})$$

$$\frac{\partial \dot{\alpha}}{\partial \dot{\alpha}_0} = 1 + \frac{2t(\dot{\delta}_0 \tan \delta_0 - \dot{r}_0)}{r_0} \quad (\text{A-12f})$$

The matrix A_3 is used to make the transformation between a spherical coordinate system and a Cartesian plane-of-sky coordinate system. It is defined as follows:

$$A_3 = \begin{pmatrix} \frac{\partial r_{c,0}}{\partial r_0} & \frac{\partial r_{c,0}}{\partial \delta_0} & \frac{\partial r_{c,0}}{\partial \alpha_0} & \frac{\partial r_{c,0}}{\partial \dot{r}_0} & \frac{\partial r_{c,0}}{\partial \dot{\delta}_0} & \frac{\partial r_{c,0}}{\partial \dot{\alpha}_0} \\ \frac{\partial r_{\delta,0}}{\partial r_0} & \frac{\partial r_{\delta,0}}{\partial \delta_0} & \frac{\partial r_{\delta,0}}{\partial \alpha_0} & \frac{\partial r_{\delta,0}}{\partial \dot{r}_0} & \frac{\partial r_{\delta,0}}{\partial \dot{\delta}_0} & \frac{\partial r_{\delta,0}}{\partial \dot{\alpha}_0} \\ \frac{\partial r_{\alpha,0}}{\partial r_0} & \frac{\partial r_{\alpha,0}}{\partial \delta_0} & \frac{\partial r_{\alpha,0}}{\partial \alpha_0} & \frac{\partial r_{\alpha,0}}{\partial \dot{r}_0} & \frac{\partial r_{\alpha,0}}{\partial \dot{\delta}_0} & \frac{\partial r_{\alpha,0}}{\partial \dot{\alpha}_0} \\ \frac{\partial \dot{r}_{c,0}}{\partial r_0} & \frac{\partial \dot{r}_{c,0}}{\partial \delta_0} & \frac{\partial \dot{r}_{c,0}}{\partial \alpha_0} & \frac{\partial \dot{r}_{c,0}}{\partial \dot{r}_0} & \frac{\partial \dot{r}_{c,0}}{\partial \dot{\delta}_0} & \frac{\partial \dot{r}_{c,0}}{\partial \dot{\alpha}_0} \\ \frac{\partial \dot{r}_{\delta,0}}{\partial r_0} & \frac{\partial \dot{r}_{\delta,0}}{\partial \delta_0} & \frac{\partial \dot{r}_{\delta,0}}{\partial \alpha_0} & \frac{\partial \dot{r}_{\delta,0}}{\partial \dot{r}_0} & \frac{\partial \dot{r}_{\delta,0}}{\partial \dot{\delta}_0} & \frac{\partial \dot{r}_{\delta,0}}{\partial \dot{\alpha}_0} \\ \frac{\partial \dot{r}_{\alpha,0}}{\partial r_0} & \frac{\partial \dot{r}_{\alpha,0}}{\partial \delta_0} & \frac{\partial \dot{r}_{\alpha,0}}{\partial \alpha_0} & \frac{\partial \dot{r}_{\alpha,0}}{\partial \dot{r}_0} & \frac{\partial \dot{r}_{\alpha,0}}{\partial \dot{\delta}_0} & \frac{\partial \dot{r}_{\alpha,0}}{\partial \dot{\alpha}_0} \end{pmatrix} \quad (A-13)$$

The nonzero entries in the matrix A_3 are as follows:

$$\frac{\partial r_{c,0}}{\partial r_0} = 1 \quad (A-14)$$

$$\frac{\partial r_{\delta,0}}{\partial \delta_0} = r_0 \quad (A-15)$$

$$\frac{\partial r_{\alpha,0}}{\partial \alpha_0} = r_0 \cos \delta_0 \quad (A-16)$$

$$\frac{\partial \dot{r}_{c,0}}{\partial \delta_0} = -r_0 \dot{\delta}_0 \quad (A-17a)$$

$$\frac{\partial \dot{r}_{c,0}}{\partial \alpha_0} = -r_0 \dot{\alpha}_0 \cos^2 \delta_0 \quad (A-17b)$$

$$\frac{\partial \dot{r}_{c,0}}{\partial \dot{r}_0} = 1 \quad (A-17c)$$

$$\frac{\partial \dot{r}_{\delta,0}}{\partial r_0} = \dot{\delta}_0 \quad (A-18a)$$

$$\frac{\partial \dot{r}_{\delta,0}}{\partial \dot{\delta}_0} = \dot{r}_0 \quad (A-18b)$$

$$\frac{\partial \dot{r}_{\delta,0}}{\partial \alpha_0} = r_0 \dot{\alpha}_0 \sin \delta_0 \cos \delta_0 \quad (A-18c)$$

$$\frac{\partial \dot{r}_{\delta,0}}{\partial \dot{\delta}_0} = r_0 \quad (A-18d)$$

$$\frac{\partial \dot{r}_{\alpha,0}}{\partial r_0} = \dot{\alpha}_0 \cos \delta_0 \quad (A-19a)$$

$$\frac{\partial \dot{r}_{\alpha,0}}{\partial \delta_0} = -r_0 \dot{\alpha}_0 \sin \delta_0 \quad (A-19b)$$

$$\frac{\partial \dot{r}_{\alpha,0}}{\partial \alpha_0} = \dot{r}_0 \cos \delta_0 - r_0 \dot{\delta}_0 \sin \delta_0 \quad (A-19c)$$

$$\frac{\partial \dot{r}_{\alpha,0}}{\partial \dot{\alpha}_0} = r_0 \cos \delta_0 \quad (A-19d)$$

N92-24306

JJ576450

P-17

Orbit-Determination Performance of Doppler Data for Interplanetary Cruise Trajectories

Part II: 8.4-GHz Performance and Data-Weighting Strategies

J. S. Ulvestad
Navigation Systems Section

A consider error covariance analysis has been performed in order to investigate the orbit-determination performance attainable using two-way (coherent) 8.4-GHz (X-band) Doppler data for two segments of the planned Mars Observer trajectory. The analysis includes the effects of the current level of calibration errors in tropospheric delay, ionospheric delay, and station locations, with particular emphasis placed on assessing the performance of several candidate elevation-dependent data-weighting functions. One weighting function has been found that yields good performance for a variety of tracking geometries. This weighting function is simple and robust; it reduces the danger of error that might exist if an analyst had to select one of several different weighting functions that are highly sensitive to the exact choice of parameters and to the tracking geometry. Orbit-determination accuracy improvements that may be obtained through the use of calibration data derived from Global Positioning System (GPS) satellites also have been investigated, and can be as much as a factor of three in some components of the spacecraft state vector. Assuming that both station-location errors and troposphere calibration errors are reduced simultaneously, the recommended data-weighting function need not be changed when GPS calibrations are incorporated in the orbit-determination process.

I. Introduction

Two-way Doppler data are the primary data type used for the navigation of robotic spacecraft during the interplanetary cruise portions of their trajectories. Empirically, it has been found that data acquired at very low elevation angles at a given tracking station can degrade the accu-

racy of the determination of a spacecraft trajectory. In the case of data taken at 2.3 GHz (S-band), the degradation occurs primarily because of errors in the calibration of the Earth's ionosphere and the interplanetary charged-particle medium, while for 8.4-GHz (X-band) data, errors in the troposphere calibration generally cause the largest Doppler measurement errors at low elevations. Therefore,

traditional practice has been to (1) impose an arbitrary elevation-angle cutoff, below which all data are discarded; or (2) apply a data-weighting function that reduces the weighting of the low-elevation data by increasing the variances assigned to Doppler measurements made at decreasing elevation angles.

Reference [1] described a consider covariance analysis methodology that was developed to investigate the effects of different Doppler weighting functions on orbit-determination accuracy for interplanetary cruise trajectories. This article reports results obtained by applying that covariance analysis to two different portions of a hypothetical trajectory for the Mars Observer mission [2], which is scheduled to be launched toward Mars in September 1992 and to arrive in September 1993. It is assumed that two-way Doppler data will be acquired at 8.4 GHz by the DSN. The calibration errors are those expected in such data during normal DSN operations in 1992 and 1993. Tracking sessions ("passes") are assumed to take place at a single Deep Space Station (either Goldstone, California, or Tidbinbilla, Australia) on 10 consecutive days. In each case, the error covariance matrix is calculated for the determination of the spacecraft state vector at an epoch defined to be the midpoint of the first of the 10 passes.

Section II of this article gives a brief review of the covariance analysis procedure developed in [1], including a definition of the data-weighting function. Section III summarizes the justification for the assumptions about the current level of calibration errors. In Section IV, a general description is given of the data arcs and assumptions employed for calculation of the error covariance matrices. Section V describes the results of the analysis for a near-zero-declination segment of the Mars Observer trajectory, while Section VI presents similar results for a high-declination trajectory segment. Section VII gives the results of an investigation of the benefits that might be obtained by reducing various calibration errors through the use of Global Positioning System (GPS) satellites or water-vapor radiometers. Section VIII provides some discussion of the results and provides a recommended weighting function for 8.4-GHz Doppler data for Mars Observer and other interplanetary spacecraft. Section IX summarizes the results and indicates some possible directions for future work.

II. Covariance Analysis and the Weighting Function

A. Introductory Remarks

The consider covariance error analysis used to investigate the sample trajectories was described in [1], and the

reader is referred to that article for details. (Also, see [3].) For each data arc, four parameters were extracted from the final covariance matrix in order to characterize the quality of the estimated spacecraft trajectory. Those four were the uncertainty in the line-of-sight position and velocity of the spacecraft as well as the lengths of the major axes of the position and velocity error ellipses in the plane of the sky, that plane perpendicular to the Earth-spacecraft line of sight. These errors are expressed in a Cartesian plane-of-sky frame rather than a spherical-coordinate frame. Thus the error in the line-of-sight velocity appears greater than it would in a spherical coordinate system because of the uncertainty in the direction of the line-of-sight coordinate axis.

B. Weighting Function

The weighting function used for a particular Doppler point has been given the form

$$w = \left[\sigma_D^2 + \left(\frac{\sigma_e}{(\sin \gamma)^q} \right)^2 \right]^{-1} \quad (1)$$

for elevation angles γ above a user-selected cutoff γ_{\min} , and zero at lower elevations (i.e., the data at elevations below γ_{\min} were discarded). In Eq. (1), σ_D is the assumed Doppler data noise (including any possible deweighting that is independent of elevation), γ is the elevation angle of the spacecraft, σ_e is a constant that determines the importance of the elevation-dependent part of the weighting function, and q is a parameter governing the form, or "steepness," of that dependence. The user has the ability to vary four different parameters (σ_D , σ_e , γ_{\min} , and q) in order to optimize results. Of course, σ_D must be no smaller than the precision of the Doppler measurements, while γ_{\min} must be no smaller than the lowest elevation at which the DSN antennas can operate.

The chosen form of the weighting function is based on the behavior of the static troposphere calibration error as it varies with elevation, and to a lesser extent, the behavior of ionosphere calibration errors as well. The weighting function for various choices of parameters in Eq. (1) was plotted in [1], as was the shape of the partial derivatives of the Doppler measurements with respect to errors in the media calibration. In [1], it was shown that troposphere calibration error was the dominant media effect at 8.4 GHz for low-elevation angles.

III. Assumed Error Magnitudes

The level of random measurement noise in the Doppler data and the magnitude of the calibration errors must be

assumed in order to estimate the errors in the spacecraft state vector at the reference time. The one-sigma noise on the Doppler data was taken to be 0.1 mm/sec over an assumed integration time of 60 seconds, typical of the data noise seen in 8.4-GHz, two-way Doppler data acquired from the Magellan spacecraft.¹

The one-sigma errors in the calibration of the zenith path delay due to the wet and dry troposphere were taken to be 4 cm and 1 cm, respectively. These errors are typical of those obtained from current calibration methods using measurements of atmospheric pressure, temperature, and relative humidity. There is some potential for reducing the error in the wet component, which tends to be the dominant error source, through the use of water-vapor radiometers [4] or GPS satellites [5].

The partial derivatives of Doppler measurements with respect to ionosphere calibration errors can be written in terms of two coefficients that represent (approximately) the uncertainties in the daytime and nighttime delays at the zenith. The calibration errors were taken to be 4 cm for the daytime coefficient and 2 cm for the nighttime coefficient, the approximate levels achievable using the current Faraday rotation measurements at the DSN Deep Space Stations [6]. At 2.3 GHz, the uncertainty in these coefficients would be a factor of ~ 13 higher because of the dependence on the square of the wavelength. GPS observations could improve these calibrations in the future [7].

The final calibration error considered is the location of the tracking station. This error source also can represent to some degree the error in the calibration of Universal Time and Polar Motion, since the partial derivatives are similar. An error in timing is equivalent to an error in the station longitude, while an error in the position of the Earth's pole will effectively manifest itself as an error in the tracking-station position relative to the pole. In the calculations presented below, the errors in the equatorial components of the station locations were assumed to be 30 cm, while the error in the z-component (toward the pole) relative to the Earth's center was taken to be 5 m. The assumed errors in the equatorial plane are slightly conservative for station locations alone, but are reasonable values when current Earth-orientation errors also are included. As with the other calibration errors, GPS observations conceivably could provide substantial reductions in uncertainty [8,9]. All the calibration errors given here

are assumed in Sections V and VI; they are summarized in Table 1.

IV. Data Arcs

All the analyses described in this article are based on one of two segments of a hypothetical Mars Observer trajectory. One segment begins at a declination near 20 deg, while the other starts near 5 deg declination. These two cases were studied because the well-known lack of sensitivity of Doppler data to the spacecraft declination for trajectory segments near the celestial equator (e.g., [10]) could lead to different conclusions for the high- and near-zero-declination cases.

For each trajectory segment, the error covariance matrices were calculated for data arcs consisting of 10 consecutive days of two-way (coherent) 8.4-GHz Doppler data acquired using a single tracking station. The reference epoch for each data arc was taken to be the midpoint of the tracking pass on the first day of that arc. For each data arc, cases were computed for low-elevation cutoffs ranging from 6 deg to 40 deg; all ten passes within a data arc were assumed to have the same cutoff. Contiguous 60-second Doppler data points were assumed to be acquired for the entire length of each pass; the magnitude of the data noise and the assumed calibration errors were as described in Section III. Combination of data from more than one tracking station goes beyond the scope of this work and was not considered. However, the cases of tracking from a Northern Hemisphere site (Goldstone, California, at a latitude of approximately +35 deg) and from a Southern Hemisphere site (Tidbinbilla, Australia, at a latitude of approximately -35 deg) were both considered. The coordinates assumed for the tracking stations are given in Table 2.²

V. Near-Zero-Declination Mars Observer Trajectory Segment

The near-zero-declination segment of the Mars Observer trajectory begins on Day 203 (22 July) of 1993. The initial geocentric state vector, in spherical coordinates, is given in Table 3. The initial spacecraft declination is 5.1 deg, but the initial declination rate of approximately -0.22 deg/day implies that the spacecraft declination is near 3.1 deg by the end of the 10-day data arc.

¹ G. R. Kronschnabl, "Magellan Differenced Doppler Data Quality Evaluation," Interoffice Memorandum 3140-GRK-90-011 (internal document), Jet Propulsion Laboratory, Pasadena, California, May 24, 1990.

² T. D. Moyer, "Station Location Sets Referred to the Radio Frame," Interoffice Memorandum 314.5-1334 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 24, 1989.

A. Tracking From Goldstone

First, consider the case in which Doppler data are acquired only from Goldstone. Although the bulk of the results in this article will be given only in figures, inspection of the results in a tabular form also can be instructive. Therefore, the computed uncertainties for the line-of-sight spacecraft position for a variety of weighting functions are compiled in Table 4. Eleven different cases are included in this table, each having the same initial spacecraft state vector. Eight different elevation-angle cutoffs were considered for each case. The first row of numbers gives the results for a covariance analysis including only the data-noise component of the measurement errors and having uniform data weighting. The remaining ten cases all include the calibration and random measurement error levels listed in Table 1.

As expected, the result for the case in which only data noise is included improves as the elevation cutoff is reduced; i.e., taking more data of uniform quality always helps, hardly a surprising result. However, if the same uniform weighting is applied to data in the presence of unmodelled media calibration errors, the large systematic Doppler measurement errors at low elevations make the result substantially worse when all the data are included (cf. second row of Table 4). The data analyst's first and most obvious recourse would be to discard the low-elevation data. An alternative of uniformly deweighting all of the Doppler data flattens the shape of the dependence on elevation cutoff, but only at the expense of making the results worse for all values of the cutoff. The third row of Table 4 shows the results for uniform deweighting from 0.1 mm/sec to 0.2 mm/sec. In fact, the data would have to be uniformly deweighted to approximately 0.5 mm/sec to remove the bulk of the dependence on the selected elevation cutoff.

Of course, the point of this investigation is to explore the benefits of nonuniform data deweighting. Results for eight choices of parameters in the weighting function (Eq. [1]) are given in the last eight rows of Table 4. As might be expected, when the data weighting is not strongly dependent on elevation angle ($q = 1$), the low-elevation data still corrupt the line-of-sight position estimates significantly. Although the results for a high-elevation cutoff can be as good as those for the steeper weighting functions, achievement of the best results still depends greatly on the exact choice of the elevation cutoff. This lack of robustness leads to a strong preference for lower weighting of the low-elevation data. In fact, for a weighting function having $q = 3$, the expected errors are almost independent of the elevation cutoff and also change little when the coefficient σ_e is varied by a factor of five between 0.01 mm/sec

and 0.05 mm/sec. The lack of dependence on the selection of the elevation cutoff occurs because the steep weighting function already has the effect of almost completely discarding the data at the lower elevations, making the specification of the cutoff superfluous.

Figures 1(a-d) display results for the four figures of merit generally used in this analysis, radial position and velocity errors, and the semimajor axes of the position and velocity error ellipses in the plane of the sky. In each figure, six curves are plotted. Curve A is the predicted result for perfect calibration of the troposphere, ionosphere, and station location. Curve B assumes standard calibration errors and uniform data weighting for a noise level of $\sigma_D = 0.1$ mm/sec. Curves C, D, and E all are based on standard calibration errors and $\sigma_e = 0.03$ mm/sec. Curve C has $q = 1$ in Eq. (1), curve D has $q = 2$, and curve E has $q = 3$. Finally, curve F corresponds to uniform weighting with the Doppler data deweighted to $\sigma_D = 0.2$ mm/sec.

In Fig. 1, the shapes of the curves as functions of weighting function and elevation cutoff angle look similar for both line-of-sight coordinates and for the plane-of-sky position. One can find differences in detail, but the general conclusions are similar for all three quantities. As expected for the near-zero-declination case, the error in the plane-of-sky position is predominantly in the declination direction, with its major axis within 15 deg of the declination axis in all cases.

The result for the plane-of-sky velocity error [Fig. 1(d)] has a different character. Even in the absence of calibration errors, this quantity is relatively poorly determined by the Doppler measurements over a short data arc, often a factor of more than 30 worse than the line-of-sight velocity. This occurs because the spacecraft plane-of-sky motion must be inferred indirectly from the spacecraft's Doppler signature over the course of the data arc, while the epoch line-of-sight velocity is measured much more directly, since the Doppler shift effectively is equivalent to the station-spacecraft range rate. The media and station-location calibration errors have little impact on the overall accuracy of determination of the plane-of-sky velocity, and deweighting or discarding low-elevation data causes a loss of information that would help provide a more accurate result. The best result for the plane-of-sky velocity is achieved in this case by choosing $q = 1$ and using data acquired at the lowest possible elevation angles.

B. Tracking From Australia

The results for tracking data from the Tidbinbilla, Australia, DSN site are quite similar to the results for tracking

data from Goldstone. In general, the expected errors are 5 to 10 percent higher for the Australia data, with dependences on the weighting functions that are virtually identical to those in the Goldstone case. Because of the similarity to the results for Goldstone, no plots are shown for the case of tracking from Australia. The slight decrease in accuracy is due to the shorter tracking passes from the southern DSN site for a spacecraft that is a few degrees north of the equator; the difference in tracking pass length ranges from 1 to 1.5 hours, depending on the elevation cutoff selected.

VI. High-Declination Mars Observer Trajectory Segment

The high-declination segment of the Mars Observer trajectory begins on Day 113 (23 April) of 1993, with an initial spacecraft declination of about 20.3 deg. Initial conditions for the spacecraft state are given in Table 5. Over the course of the 10-day data arc, the spacecraft declination decreases to about 19.5 deg.

A. Tracking From Goldstone

Figures 2(a-d) display the results for the 4 standard figures of merit in the case of 10 consecutive tracking passes from Goldstone. Predicted orbit-determination accuracy is a factor of 2 to 8 better than that for the near-zero-declination trajectory segment, depending on the specific component and the weighting function being investigated. (Note that the spacecraft distance from Earth is only 60 percent of its distance during the near-zero-declination trajectory segment.) A notable difference [Fig. 2(a)] is that the line-of-sight position error is quite close to the results for data noise only, implying that the current calibration errors have relatively little effect on the determination of this quantity using two-way Doppler data. For this case, the weighting functions with little deweighting at low elevations (i.e., $q = 1$ rather than 2 or 3) give the best results, since the power of the extra data is more important than the small increase in errors due to miscalibration. However, results for $q = 2$ are close to the results for $q = 1$ for both line-of-sight position and plane-of-sky velocity [Figs. 2(a) and 2(d)], while they are somewhat better for the other two variables [Figs. 2(b) and 2(c)]. In general, the specific choice of weighting functions is much less important than it is for the near-zero-declination case. Some deweighting of the low-elevation data still is important, however, in order to prevent sharp upturns in the trajectory uncertainty plots when data below 10 deg elevation are included in the orbit solutions.

B. Tracking From Australia

One might expect the tracking from Australia to give somewhat poorer results than that from Goldstone because of the abbreviated tracking passes (8.8 hours instead of 12.9 hours for tracking down to 6 deg elevation) and because of the larger amount of low-elevation data. While this is generally true, the degradation in performance is not very great. In fact, the Australia data give performance a factor only 1.2 to 2 poorer than the Goldstone data for all figures of merit. When the line-of-sight position error is considered for the cases in which there is little deweighting, the performance degradation factor of ~ 1.2 is roughly the square root of the difference in tracking-pass lengths, as expected for the situation in which the results are dominated by data noise.

Figures 3(a-d) show results for several different weighting functions. The best results for both line-of-sight position and plane-of-sky velocity [Figs. 3(a) and 3(d)] for the Australia track are achieved by applying no data deweighting and using data acquired at elevation angles as low as 6 deg. Some deweighting is useful for the line-of-sight velocity [Fig. 3(b)], and it is especially important for the plane-of-sky position [Fig. 3(c)]. Results for $q = 2$ appear to be best and are 20 to 40 percent better than those for the case with $q = 1$; the cases having $q = 2$ are about 20 to 40 percent *worse* than those for $q = 1$ when the line-of-sight position and plane-of-sky velocity are considered. Deweighting the low-elevation data too much actually degrades the accuracy of the latter two components of the state vector because they depend more on data noise than on the accuracy of the calibration of systematic errors.

VII. Effects of Improved Calibration

It is well-known that measurements derived from the GPS satellites have the potential for improving the calibration of some of the effects that lead to errors in Doppler data [5,7-9]; water-vapor radiometers [4] also might offer some improvement in the calibration of the wet troposphere. The orbit-determination improvement that might be obtained has been investigated by reducing the levels of the troposphere calibration errors and the station-location errors, both individually and together, in the analyses described in Sections V and VI. Reductions in ionosphere calibration errors have not been considered because they have less effect on 8.4-GHz data. The calibration improvements that were considered were a decrease from 4 cm to 1 cm in the calibration error for the zenith delay due to the wet troposphere, and an improvement to 5 cm in the knowledge of each of the three station-location components.

Doppler tracking from Goldstone with improved calibration of the systematic errors has been investigated for both the near-zero- and high-declination cases. (Results for tracking from Australia would be similar.) For the near-zero-declination trajectory segment, Figs. 4(a-d) show a sample of results for cases in which the wet-troposphere and station-location calibrations were improved by the amounts specified above. Curve A is the comparison curve for uniform weighting with $\sigma_D = 0.1$ mm/sec, derived under the assumption of perfect calibration. Curve B shows the results for a weighting function having $\sigma_e = 0.03$ mm/sec, $q = 2$, and $\sigma_D = 0.1$ mm/sec, assuming the current calibration errors given in Table 1. Curve C shows the results for the same weighting function with the zenith wet troposphere calibration error reduced from 4 cm to 1 cm, curve D shows results with the station-location calibration errors reduced to 5 cm per component, and curve E shows the results when both the station-location and troposphere calibrations are improved simultaneously.

It is apparent that improving the station-location calibration to 5 cm per component provides a greater error reduction than does the improvement from 4 cm to 1 cm in the calibration of the wet troposphere. A more detailed investigation would be necessary to assess the value of incremental improvements, say, improving the wet troposphere calibration from 4 cm to 3 cm or the z -component of the station-location error from 5 m to 50 cm. Improving both calibrations to levels that may be achievable using observations of GPS satellites can reduce the orbit-determination errors by as much as a factor of three for some parameters.

Figures 4(a-c) show that when only the station location calibrations are improved, a weighting function having $q = 2$ actually does not completely flatten the dependence on the elevation-angle cutoff. This clearly is because the troposphere calibration error dominates the orbit-determination error for this case, and the low-elevation data should have even lower weights. Although the weighting function with $q = 3$ (not shown) actually would be a better choice when only the station-location errors are reduced, $q = 2$ is the preferred choice when both station-location and troposphere calibration errors are reduced.

Investigation of the high-declination trajectory segment shows similar improvements from reducing the calibration errors, as displayed in Figs. 5(a-d). Again, the improvement in the station-location calibration yields the biggest gain. Naturally, the improvements are less in the parameters already dominated by data noise rather than by calibration errors. But improvement by more than a factor of three still is possible for the plane-of-sky position.

VIII. Discussion

The figures and tables presented in the preceding sections showed results for a variety of assumptions. However, for the current level of calibration errors, the general character of the results did not change significantly over the different cases considered. In particular, inspection of Figs. 1-3 shows that the curves labelled D generally have the best combination of attributes for the line-of-sight and plane-of-sky position and velocity uncertainties. For low-elevation cutoffs, these curves show a significant reduction of errors over the uniformly weighted case. They also appear quite flat as a function of the selected elevation cutoff. Therefore, the weighting function labelled D appears to be a good choice for 8.4-GHz Doppler data with noise of 0.1 mm/sec that is affected by calibration errors at the levels assumed herein. The parameters of this weighting function are $\sigma_D = 0.1$ mm/sec, $\sigma_e = 0.03$ mm/sec, and $q = 2$.

For the near-zero-declination trajectory and tracking from Goldstone, Figs. 6(a) and 6(b) show the respective dependence of predicted position and velocity accuracies on σ_e , assuming a low-elevation cutoff of 6 deg, $q = 2$, and $\sigma_D = 0.1$ mm/sec. Note that the predicted errors in both position components and in the line-of-sight velocity actually are smallest for $\sigma_e = 0.05$ mm/sec, but vary by less than ~10 percent for σ_e between 0.02 and 0.05 mm/sec. The plane-of-sky velocity, on the other hand, is dominated by the data noise and is determined most accurately for $\sigma_e = 0.01$ mm/sec. The choice of $\sigma_e = 0.03$ mm/sec appears to be a good compromise. If the plane-of-sky velocity is not a parameter of great interest, $\sigma_e = 0.05$ mm/sec would be a slightly better choice. The trajectory solution statistics seem fairly robust for choices of σ_e between 0.02 and 0.05 mm/sec; the best selection for real orbit-determination analysis should be established by making tests with actual spacecraft data rather than relying solely on a covariance analysis.

Based on the results of the analysis, the following weighting function is recommended for 8.4-GHz two-way Doppler data acquired from Mars Observer:

$$w = \left[(0.1 \text{ mm/sec})^2 + \left(\frac{0.03 \text{ mm/sec}}{\sin^2 \gamma} \right)^2 \right]^{-1} \quad (2)$$

where γ is the elevation angle. All data down to the lowest possible elevation angle should be included in orbit-determination calculations. The above weighting function also is recommended for the interplanetary cruise trajectory of any other mission having calibration errors

similar to those expected for Mars Observer. The recommendation holds for trajectories at different declinations and for tracking from any of the three main DSN complexes (California, Australia, or Spain). Operational orbit-determination software should incorporate a weighting function of the above form, with some capability for analysts to vary the parameters slightly as necessary to achieve the best estimated trajectory.

Analysts processing the actual data for missions such as Mars Observer might find slight improvements in the results by varying the parameters in Eq. (2). For instance, they may find that the data below 10-deg elevation angles seem to corrupt their solutions, possibly because the effects of the troposphere calibration errors have been underestimated or because the time-varying component of the troposphere (not considered in the work presented here) affects the data. A somewhat larger value of σ_e or a cutoff in elevation angle may improve the results under those circumstances. Of course, given that other data types such as ranging and Very Long Baseline Interferometry (VLBI) sometimes are included in the orbit-determination process, the best weighting function for 8.4-GHz Doppler data when multiple data types are used could be slightly different from that shown in Eq. (2).

Perhaps the most significant attribute of the weighting function recommended here is that it removes almost all dependence on the selection of an elevation-angle cutoff. In fact, the magnitudes of the predicted errors are quite similar to the errors expected for uniformly weighted data if the proper elevation cutoff is selected. However, the recommended solution is more robust, virtually removing the possibility of corrupting the estimated trajectory due to a poor choice of the elevation cutoff. In fact, in the case of uniformly weighted data, a much lower elevation-angle cutoff generally would be preferred for determining the plane-of-sky velocity than for the other orbit parameters, making it quite difficult to choose the best compromise. In addition, dependence of the preferred weighting function on the identity of the station is eliminated when Eq. (2) is used as the weighting function. If uniform weighting were used for a high-declination trajectory, the preferred elevation cutoff for tracking from Australia would be considerably lower than that for tracking from California, and there would be a risk that the wrong choice of cutoff angles might be made for the different stations.

Elimination of the elevation-angle cutoff also is beneficial from a navigation standpoint because the inclusion of all possible data yields longer effective passes from each station. Although only the interplanetary cruise phase has been analyzed, it is likely that a weighting function simi-

lar to Eq. (2) would be helpful in planetary approach and encounter scenarios, when trajectory solutions sometimes must be generated from only a few passes of data. Longer passes would provide more sensitivity to spacecraft acceleration caused by a planet's gravitational field.

For the cases in which improved calibrations are assumed, the weighting function shown in Eq. (2) also seems to give good results. Greater deweighting of the elevation data ($q = 3$) might be preferable for the case in which only the station-location calibration is improved. However, if it is assumed that the troposphere and station-location calibration errors are reduced at approximately the same time (for instance, both by using GPS-based calibration data), a weighting function similar to that specified in Eq. (2) still seems to be a good choice.

IX. Summary and Directions for Future Work

This article has reported the results of a covariance analysis of the use of 8.4-GHz Doppler data for orbit determination in the interplanetary cruise phase of a mission. Investigation of a variety of data-weighting functions has resulted in selection of an elevation-dependent data-weighting scheme. Predicted accuracies are similar to those that conceivably could be achieved with uniformly weighted data and a cutoff at high-elevation angles. However, use of the recommended weighting function and inclusion of data taken at all elevations enhances the robustness of the orbit-determination process. The same weighting function will work well for high- and near-zero-declination data and for all DSN tracking stations. Thus the recommended weighting function greatly reduces the sensitivity of the estimated trajectory to choices left to the data analyst's discretion.

The results presented here represent a simplified case that still should be applicable to the more complex circumstances encountered in operational orbit determination. However, specific analyses should be performed for more complicated situations. The list below suggests several areas in which future work would be beneficial.

- (1) Analyze the results obtainable for a 10-day data arc including Doppler data taken from all three DSN sites, not just one of those sites.
- (2) Analyze the accuracy obtainable at 32 GHz, with lower data noise and a variety of assumptions about the calibration errors.
- (3) Extend the analysis method to include the gravitational influence of another planet, enabling applica-

tion of the calculations to the planetary approach case.

- (4) Perform a similar analysis for other data types, such as range and VLBI.
- (5) Combine the different data types to find an optimal weighting for each in the presence of the other data in the orbit-determination solutions.

The analysis suggested in items (1) and (2) above represents a fairly straightforward extension of this article and [1]. However, items (3–5) all require significant increases in model and computational complexity. Thus, they would make it more difficult to consider a large number of different cases in a short period. The desirability of spending the effort necessary to make those investigations has not yet been established.

Acknowledgments

The author thanks Stuart Demczak for supplying the Mars Observer trajectory used for the studies in this article, and Sam Thurman for many helpful discussions and critical readings of earlier drafts of the article.

References

- [1] J. S. Ulvestad and S. W. Thurman, "Orbit-Determination Performance of Doppler Data for Interplanetary Cruise Trajectories, Part I: Error Analysis Methodology," *TDA Progress Report 42-108*, vol. October–December 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 31–48, February 15, 1992.
- [2] M. Mecham, "Mars Observer Begins New Era Using Proven Spacecraft Design," *Aviation Week and Space Technology*, vol. 131, no. 15, pp. 79–82, October 9, 1989.
- [3] V. J. Ondrasik and D. W. Curkendall, "A First-Order Theory for Use in Investigating the Information Content Contained in a Few Days of Radio Tracking Data," *Deep Space Network Progress Report*, JPL Technical Report 32-1526, vol. III, pp. 77–93, June 15, 1971.
- [4] J. M. Moran and B. R. Rosen, "Estimation of the Propagation Delay Through the Troposphere from Microwave Radiometer Data," *Radio Science*, vol. 16, no. 2, pp. 235–244, March–April, 1981.
- [5] S. M. Lichten, "Precise Estimation of Tropospheric Path Delays With GPS Techniques," *TDA Progress Report 42-100*, vol. October–December 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 1–12, February 15, 1990.
- [6] H. N. Royden, R. B. Miller, and I. A. Buennagel, "Comparison of NAVSTAR Satellite L-Band Ionospheric Calibrations with Faraday Rotation Measurements," *Radio Science*, vol. 19, no. 3, pp. 798–804, May–June 1984.
- [7] G. E. Lanyi and T. Roth, "A Comparison of Mapped and Measured Total Ionospheric Electron Content Using Global Positioning System and Beacon Satellite Observations," *Radio Science*, vol. 23, pp. 483–492, July–August 1988.
- [8] S. M. Lichten and W. I. Bertiger, "Demonstration of Sub-Meter GPS Orbit Determination and 1.5 Parts in 10^8 Three-Dimensional Baseline Accuracy," *Bulletin Geodesique*, vol. 63, no. 2, pp. 167–189, 1989.

- [9] A. P. Freedman, "Determination of Earth Orientation Using the Global Positioning System," *TDA Progress Report 42-99*, vol. July-September 1989, Jet Propulsion Laboratory, Pasadena, California, p. 1-11, November 15, 1989.
- [10] T. W. Hamilton and W. G. Melbourne, "Information Content of a Single Pass of Doppler Data from a Distant Spacecraft," *JPL Space Programs Summary 37-39*, vol. III, Jet Propulsion Laboratory, Pasadena, California, pp. 18-23, March-April 1966.

Table 1. Assumed (one-sigma) Doppler random and systematic errors.

Parameter	Value
Data noise, mm/sec	0.1
Wet zenith troposphere, cm	4
Dry zenith troposphere, cm	1
Daytime zenith ionosphere, cm	4
Nighttime zenith ionosphere, cm	2
Station spin radius, cm	30
Station longitude, cm	30
Station z-height, m	5

Table 2. Assumed tracking station coordinates.

Station/coordinate	Value
Goldstone	
Spin radius, km	5203.997
East longitude, deg	243.1105
z-height, km	3677.052
Australia	
Spin radius, km	5205.251
East longitude, deg	148.9813
z-height, km	-3674.749

Table 3. Initial state vector for near-zero-declination segment of Mars Observer trajectory.

Parameter	Value
Year	1993
Day of year	203 (22 July)
Distance, km	3.1573×10^8
Right ascension (α), deg	169.0252
Declination (δ), deg	5.1308
Radial velocity, km/sec	11.5770
$d\alpha/dt$, deg/sec	6.2454×10^{-6}
$d\delta/dt$, deg/sec	-2.5866×10^{-6}

Table 4. Line-of-sight position errors (1σ , km) for near-zero-declination trajectory.

σ_D	σ_e	q	$\gamma_{min}, \text{ deg}$							
			6	10	15	20	25	30	35	40
0.1	NOISE	...	31.8	33.1	35.0	37.4	40.5	44.8	51.0	60.7
0.1	0	...	184.8	119.3	90.7	79.3	73.9	71.1	70.8	74.9
0.2	0	...	192.8	132.3	109.1	102.4	101.9	105.3	113.2	129.1
0.1	0.02	1	131.4	105.9	87.6	78.4	73.6	71.0	70.8	74.9
0.1	0.02	2	81.7	80.7	78.3	75.2	72.6	70.8	70.8	75.0
0.1	0.02	3	72.7	72.7	72.5	72.0	71.2	70.5	70.9	75.2
0.1	0.03	1	115.9	99.2	85.5	77.7	73.4	71.0	70.8	75.0
0.1	0.03	2	78.1	77.5	76.1	74.0	72.1	70.7	70.9	75.1
0.1	0.03	3	72.2	72.2	72.1	71.8	71.3	70.9	71.5	75.7
0.1	0.01	3	75.6	75.5	75.0	73.8	72.1	70.6	70.7	74.9
0.1	0.05	3	72.7	72.7	72.6	72.5	72.2	72.0	72.7	76.8

Table 5. Initial state vector for high-declination segment of Mars Observer trajectory.

Parameter	Value
Year	1993
Day of year	113 (23 April)
Distance, km	1.9289×10^8
Right ascension (α), deg	120.1785
Declination (δ), deg	20.2819
Radial velocity, km/sec	18.4554
$d\alpha/dt$, deg/sec	6.3462×10^{-6}
$d\delta/dt$, deg/sec	-1.0732×10^{-6}

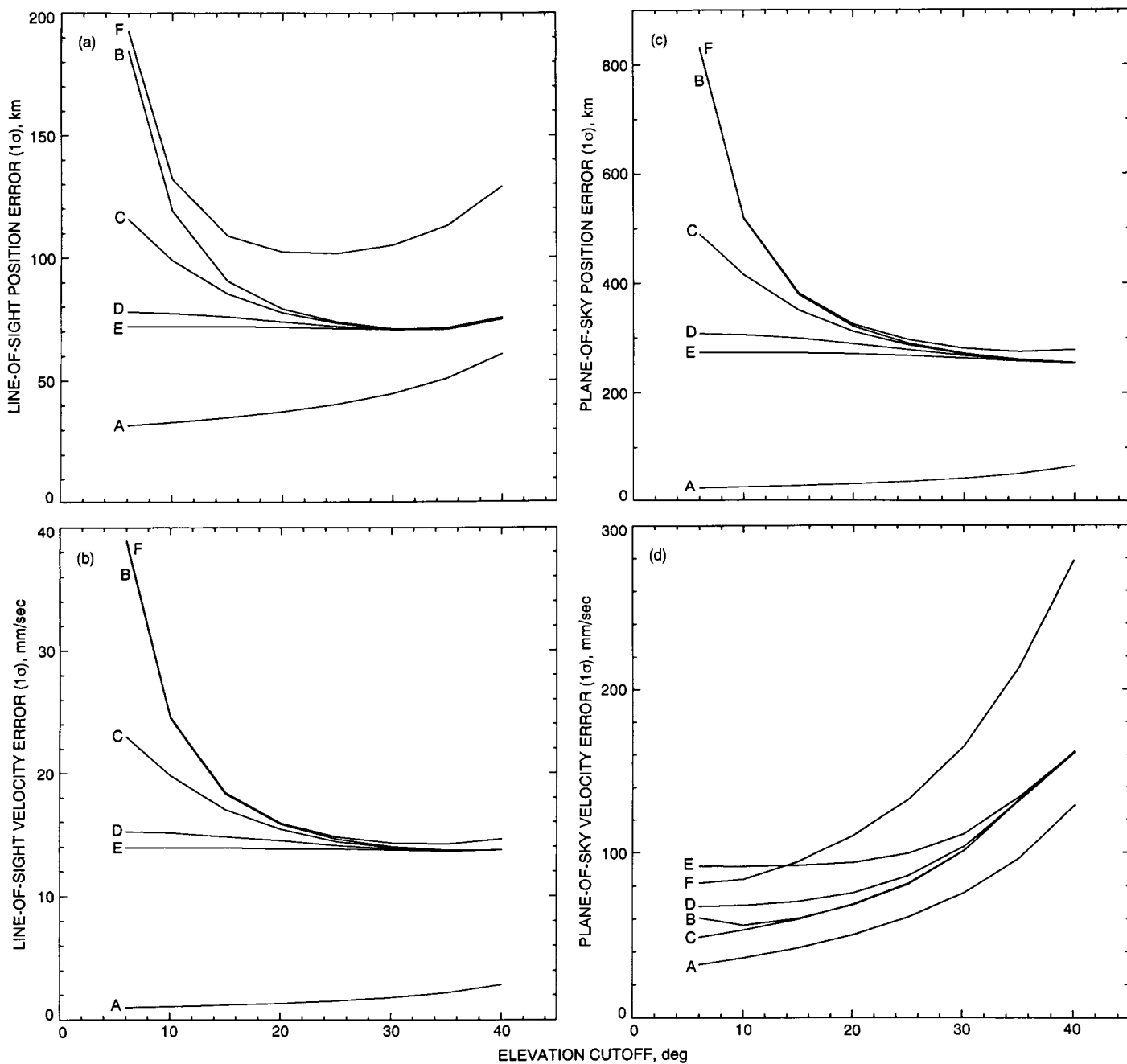


Fig. 1. Estimated errors as functions of elevation-angle cutoff for different weighting functions for the near-zero-declination segment of the Mars Observer trajectory, assuming two-way 8.4-GHz Doppler data acquired at Goldstone: (a) line-of-sight position error; (b) line-of-sight velocity error; (c) plane-of-sky position error; and (d) plane-of-sky velocity error.

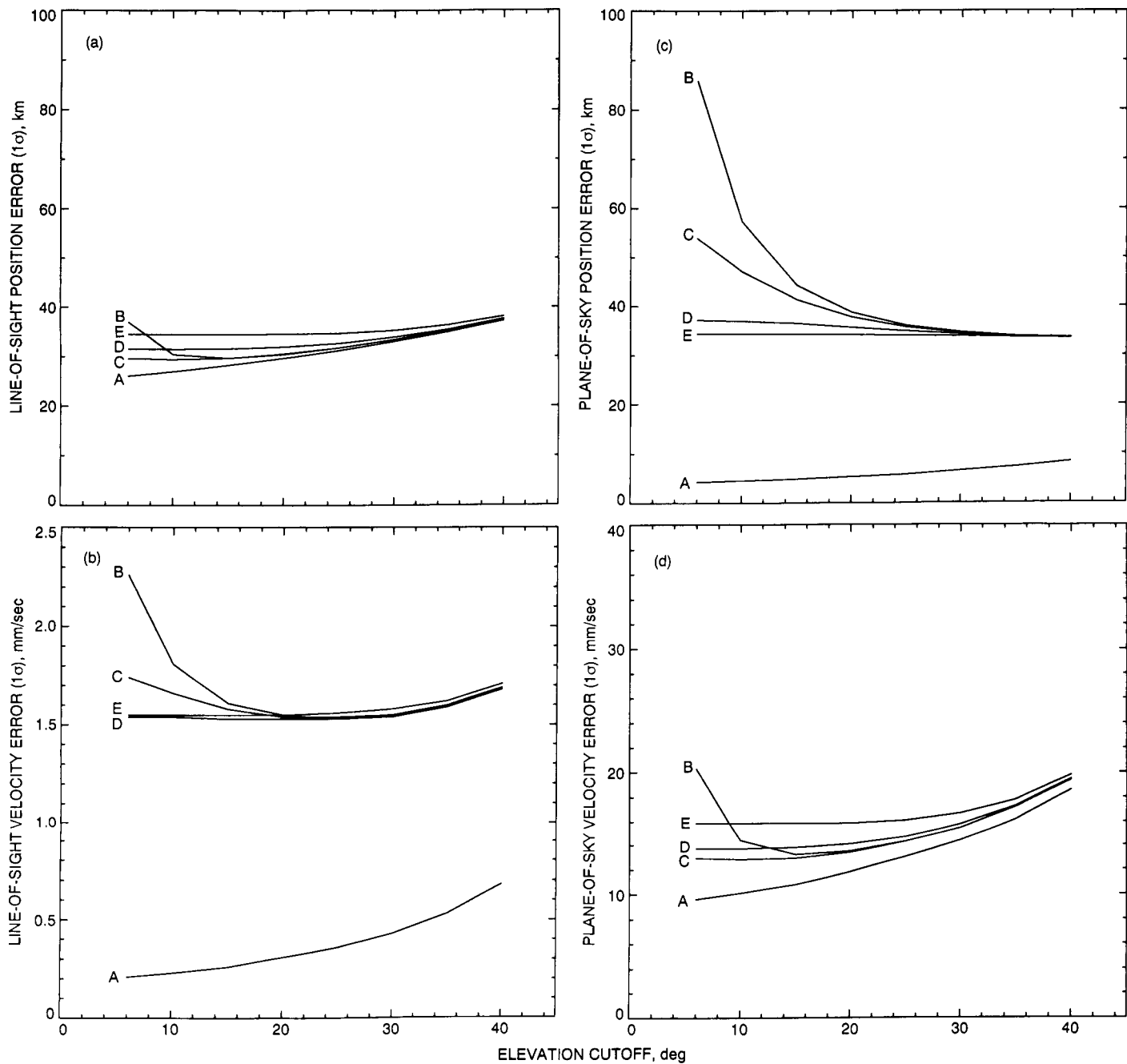


Fig. 2. Errors for the high-declination trajectory segment for Goldstone Deep Space Station. The weighting functions are the same as those used in Fig. 1, except the uniform deweighting to 0.2 mm/sec is not shown: (a) line-of-sight position error; (b) line-of-sight velocity error; (c) plane-of-sky position error; and (d) plane-of-sky velocity error.

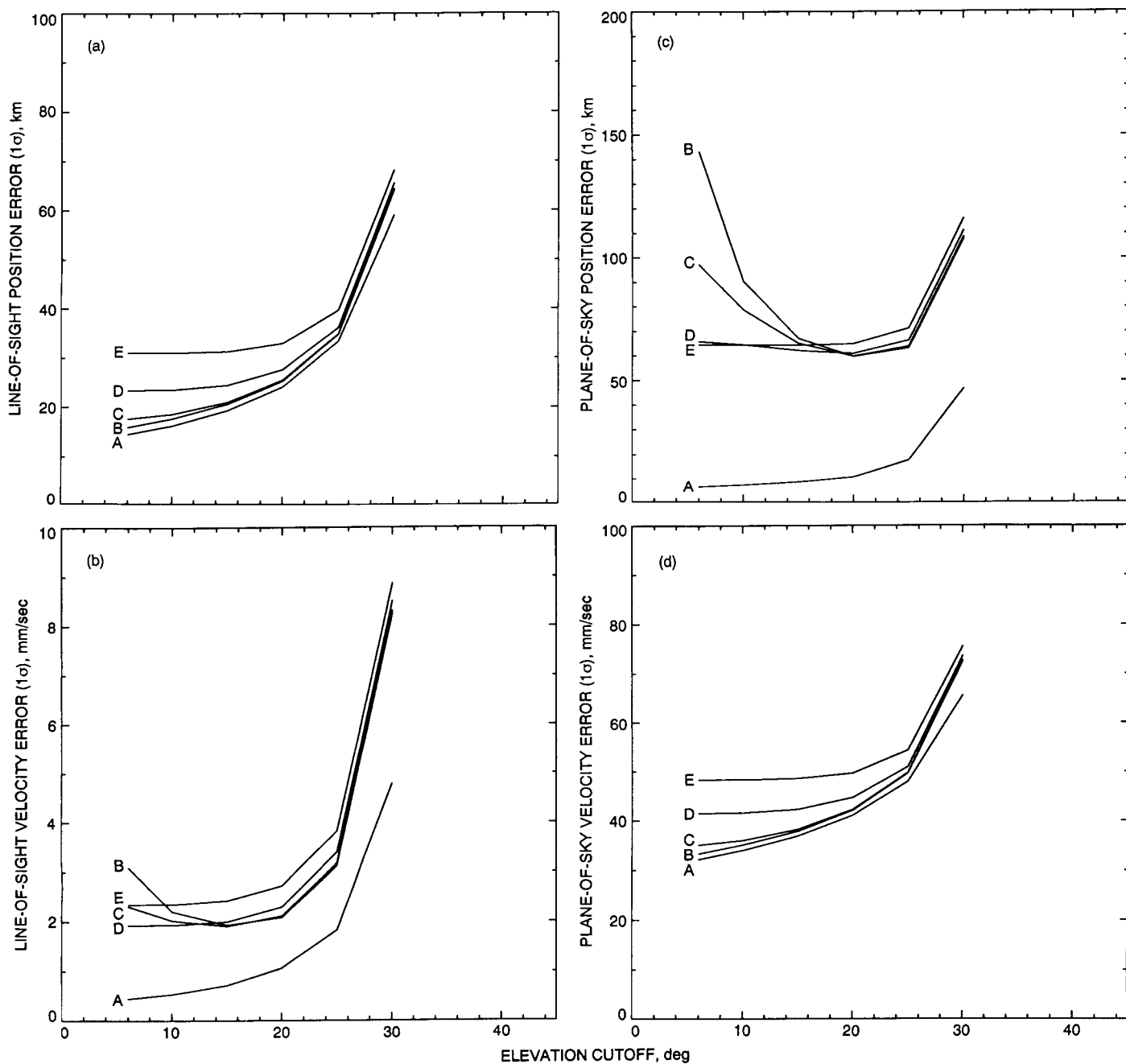


Fig. 3. Errors for the high-declination trajectory segment for Tidbinbilla Deep Space Station; weighting functions are the same as those used for Fig. 2: (a) line-of-sight position error; (b) line-of-sight velocity error; (c) plane-of-sky position error; and (d) plane-of-sky velocity error.

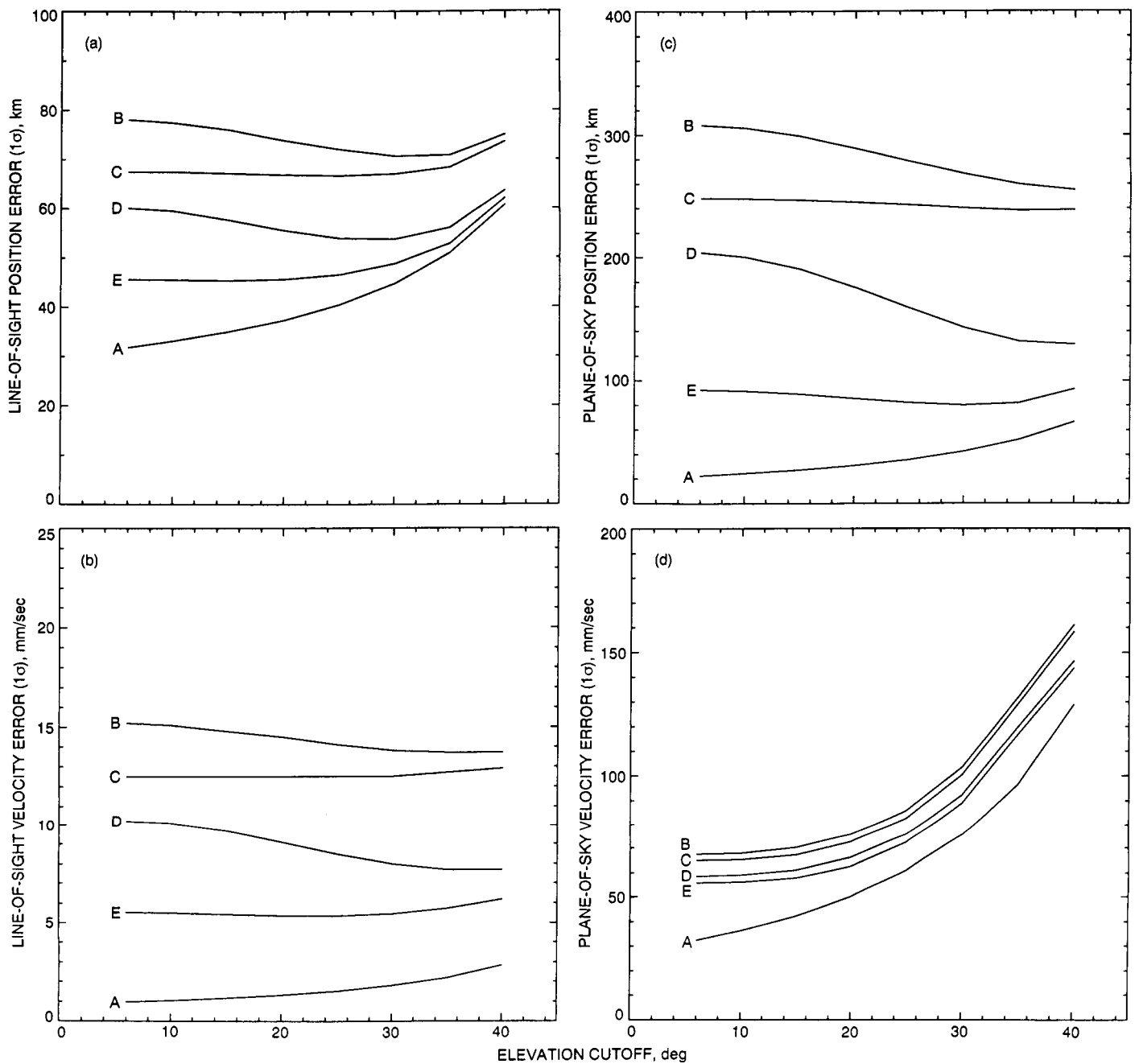


Fig. 4. Errors for the near-zero-declination Mars Observer trajectory segment, tracked from Goldstone, for different improvements in the calibration: (a) line-of-sight position error; (b) line-of-sight velocity error; (c) plane-of-sky position error; and (d) plane-of-sky velocity error.

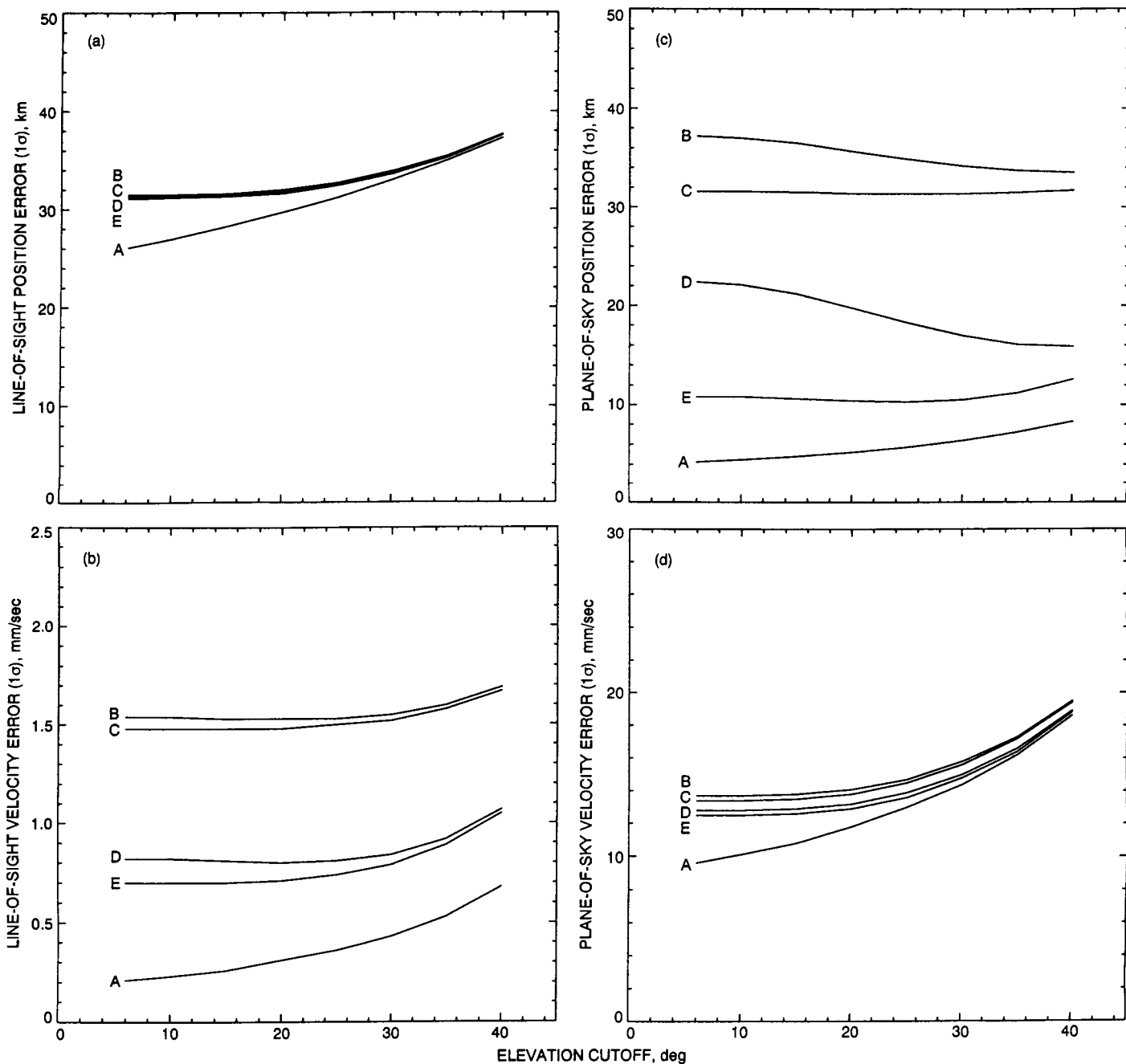


Fig. 5. Errors for the high-declination trajectory segment, tracked from Goldstone, for different improvements in the calibration: (a) line-of-sight position error; (b) line-of-sight velocity error; (c) plane-of-sky position error; and (d) plane-of-sky velocity error.

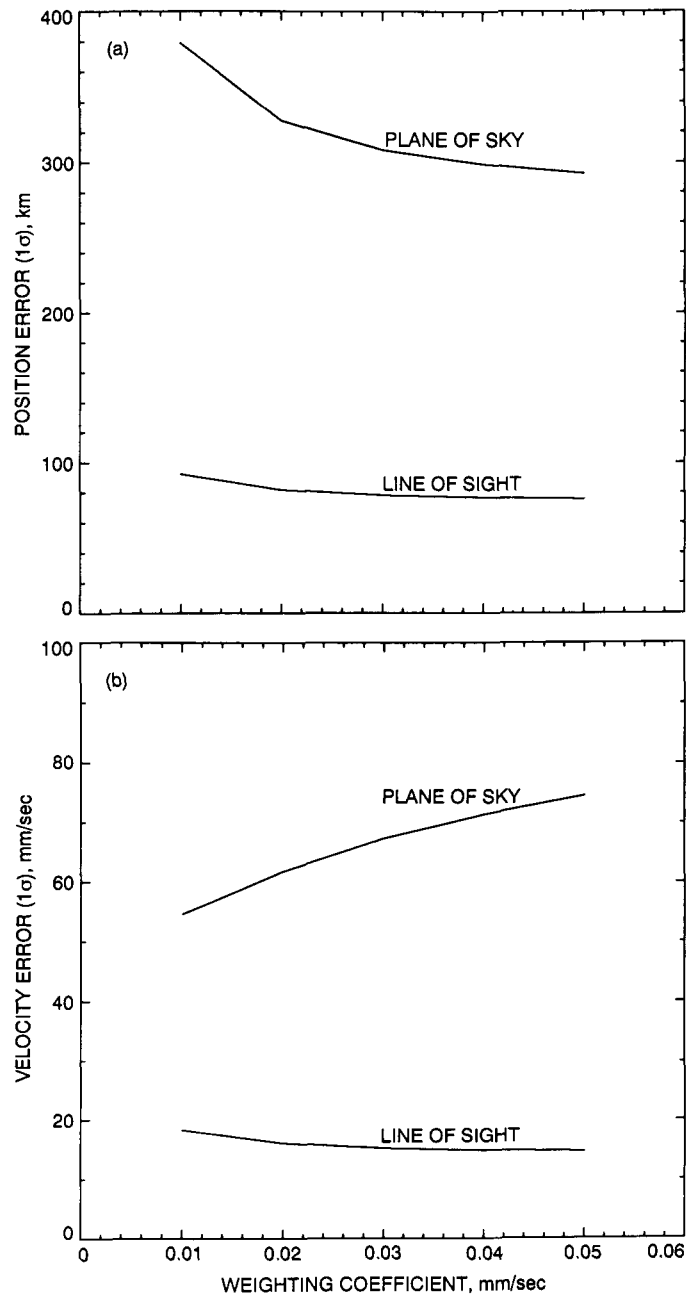


Fig. 6. The dependence of predicted position and velocity accuracies on σ_θ : (a) the dependence of the line-of-sight and plane-of-sky position errors versus σ_θ and (b) dependence of line-of-sight and plane-of-sky velocity errors versus σ_θ .

7-17
7-26
8-13

92-24307

The X-Windows Interactive Navigation Data Editor

G. C. Rinker
Navigation Systems Section

A new computer program called the X-Windows Interactive Data Editor (XIDE) has been developed and demonstrated as a prototype application for editing radio metric data in the orbit-determination process. The program runs on a variety of workstations and employs pull-down menus and graphical displays, which allow users to easily inspect and edit radio metric data in the orbit data files received from the DSN. The XIDE program is based on the OSF/Motif (Open Software Foundation) Graphical User Interface (GUI) and has proven to be an efficient tool for editing radio metric data in the navigation operations environment. It has been adopted by the Magellan Navigation Team as their primary data-editing tool. Because the software was designed from the beginning to be portable, the prototype has been successfully moved to new workstation environments. It has also been integrated into the design of the next-generation software tool for DSN multimission navigation interactive launch support.

I. Introduction

The fundamental software system used for spacecraft navigation at JPL is the Orbit Determination Program (ODP). This program, which actually consists of a sequence of integrated programs, is currently being used to support NASA deep space missions and also provides launch support for other domestic and international missions. One of the critical steps in the orbit-determination process is editing radio metric data. The X-Windows Interactive Data Editor (XIDE) is a new computer program developed as a prototype to enable navigation operations personnel to perform more efficiently in a workstation environment. Figure 1 shows an overview of the XIDE input and output relative to the ODP. The overview begins with the radio metric data that are delivered from the DSN stations through the Ground Communications Facility (GCF) to the DSN Navigation Facility, within the Network Operations Control Center (NOCC), where it is converted by

the Radio Metric Data Conditioning (RMDC) team into an orbit data file (ODFILE) for input into the ODP.

The primary function of the orbit-determination process is to estimate the spacecraft trajectory and other relevant parameters by using the processed ODFILE provided by the RMDC team. The ODP includes a complete set of algorithms necessary to model the radio metric observations and the motion of a planetary orbiter or deep space probe [1]. The program is capable of estimating the spacecraft trajectory and associated parameters by using either weighted least-squares or batch-sequential estimation algorithms.

At the heart of the estimation process, which is frequently referred to as "fitting" the data, are the residuals; for each measurement (such as Doppler or range), the residual consists of the difference of the observed value and the computed (predicted) value of the measurement. The

ODP produces a file of before-the-fit residuals, called a REGRES file, and a file of after-the-fit residuals, called a PRESID file. These residuals are the tools that navigation-operations personnel use to assess the quality of the radio metric data and the accuracy of the mathematical models used for trajectory estimation and propagation for a particular spacecraft.

It is at this stage, when the residuals are available, that the orbit-determination analyst (or radio metric data-conditioning analyst) edits the radio metric data by removing data that have unusually large measurement errors or by removing data for which the mathematical model is clearly incorrect, and then the data are fit to obtain an improved solution. The complexity of the navigation system, including measurement, data processing, and communication components makes editing radio metric data a difficult task. Success depends on the skills of the analyst, who relies on his or her experience to solve complex problems in system development and operations for navigation [2]. Navigation operations have yet to become a routine set of procedures. When the data editing has been completed by the analyst, the result is a set of delete commands that are then used in the ODP to update the orbit estimate until finally a best-trajectory estimate in the form of a PFILE is produced.

This article traces some of the significant concepts in the development of the XIDE program and shows some of the interactive features that have made it useful. However, it is neither a tutorial introduction to the subject, nor an XIDE user's guide. Illustrations of the menus and displays generated by XIDE are included, and even though their contents are not described completely, it is hoped that the reader can appreciate how the XIDE GUI helps to streamline the orbit-determination process.

II. Program Development

The availability of workstations with high-resolution bit-mapped displays and pointing devices for navigation software development, and the introduction of such hardware into the navigation-operations environment, have made it possible to exploit these capabilities. The advent of new windowing software, such as the X Window System (referred to colloquially as X-Windows) gives the developer the power to present a detailed graphical display to the user and then allows the user to interact with it.

The first operational navigation application to move in this direction is XIDE, which was developed as a prototype to demonstrate how GUI technology can improve the

productivity of navigation operations personnel. This program is the first navigation application at JPL to fully exploit the advantages of a modern workstation environment. There was a predecessor of XIDE, called IDE (Interactive Data Editor), which identified the basic concept; however, the GUI development for the workstation environment is completely new and has evolved as a cooperative effort between programmers and users.

One of the underlying goals in applying the GUI technology to a navigation application was to select a software-development strategy that would provide software portability and vendor independence. The strategy selected was based on the X-Window System, which is a network-based graphics windowing system. It was developed at the Massachusetts Institute of Technology (MIT) and has been adopted as an industry standard [3]. X-Windows provides the bare bones of a windowing system upon which almost any style of GUI can be built. One of the most popular user interfaces currently available is the OSF/Motif GUI, developed by the Open Software Foundation [4], which was adopted for software development to maximize portability.

The primary design goal for XIDE was ease of use. Based on demonstrations of XIDE, the learning curve for this program was found generally to be small, and almost nonexistent for users who have had to edit radio metric data and have used a computer with a GUI, such as a Macintosh. Users report that it takes about fifteen minutes to learn to operate the basic features of XIDE. Since all the options are represented graphically on-screen, almost no keyboard input is required (although advanced users can use it as a shortcut) because users can point and click with a mouse at the functions to be performed. To delete a data point, users simply point at the one they want to eliminate or draw a lasso around groups they want to erase. If the user wants to change the numeric scale that is used to display the data, there are a variety of ways to do this easily.

The other major design goal for XIDE was speed. Because the program is very interactive, users immediately see changes in the residual plot as they delete the data on-screen. They can iteratively select, delete, and undo deletes until the data set is satisfactorily edited. XIDE will then quickly write out a set of accurate delete commands that are executed by an ODP link when desired.

Users of XIDE have developed a second unexpected use of the application. Instead of being used solely to edit data, the program is often used simply to view and print residual plots. With XIDE's scaling and labeling capabilities, users can get a plot to appear with the desired

content on-screen before having to print it out. Then all a user needs to do is use the button labeled Print to get a hard copy.

III. Program Usage

This section gives an overview of the functions that XIDE can perform. The functions will be described in terms of an operational scenario, the basic GUI features, and the advanced GUI features.

A. Operational Scenario

A walk-through of a basic data-editing session follows: The user first sees the top-level window, as shown in Fig. 1. This consists of a menu-bar at the top, an interactive plotting area in the middle, and a message area at the bottom. By dragging the mouse, with the left button down, across the menu-bar, each of the main pull-down menus is displayed to give the user a quick overview of the program capability and organization. The menus that are displayed in this fashion are shown in Figs. 2, 3, 4, 5, and 6 and are discussed in more detail below.

The next step is to pull down the File menu, which actually appears as shown in Fig. 7. One of the features that makes this type of GUI easy to learn is that menu items that are not reasonable at this stage are grayed-out. This means that they are both visually and functionally disabled as a choice for the user. As shown in Fig. 7, at this stage all the File menu items are grayed-out except for the Open and Exit selections, since they are the only reasonable choices. On the right-hand side of the menu are the optional keyboard accelerators; for example, the keyboard combination Meta-O will select the Open item, if the user prefers to use this rather than mouse input.

In this case, when the user selects the Open item, a dialog box pops up, as shown in Fig. 8. This file-selection dialog box allows the user to move through the file-directory tree to identify the specific REGRES or PRESID file from the ODP that is to be opened for editing. Once the data file is selected, and if it is of the acceptable file type, then the Open New File dialog box is immediately replaced with a Data Set dialog box, as shown in Fig. 9. The user is presented with a complete and detailed list of the available data and can then select with the mouse the data set to be plotted from those available on the specified input file. A residual plot is then immediately displayed. After using the mouse to identify some "bad" points and then using the pull-down menus to Cut and then to Center Scale, the resulting plot of the residual data is displayed as shown in Fig. 10. The horizontal scale is time, labeled for

year, month, day, hour, minute, and second, and the vertical scale is residual value, labeled for the data type and data band (uplink/downlink). The data shown in Fig. 10 are for the Magellan spacecraft, specifically the two-way S-band Doppler data before-the-fit from August 15, 1990 at 09:00:00 UTC to August 17, 1990 at 07:00:00 UTC.

One can then edit other data types, such as the two-way minus three-way Doppler data, and if ODP delete commands are desired for this session, then the Save As command under the File menu can be selected. The resulting pop-up dialog box is shown in Fig. 11. One noteworthy item is that the delete-command format can be used for the ODP or the Orbit Data Editor (ODE). The ODE format is provided so that this tool, or one of its offspring, could be used by the DSN RMDC team and would potentially benefit all flight projects in the pre-ODP data-conditioning phase. After the user specifies the delete-command information (a file name is sufficient), then the delete commands are written in an intelligently compressed format to a text file, as shown in Fig. 12. These delete commands can then be used to update the spacecraft-orbit estimate with the ODP and finally generate an improved trajectory estimate, as shown in Fig. 1.

B. Basic GUI Features

The basic GUI features are accessed through the menu-bar, as shown in Fig. 1. The menu-bar consists of five pull-down menus: File, Edit, View, Options, and Help. The File menu (Fig. 2) allows the user to perform global input and output operations. This menu provides the capability to open a new data set (for the current input file), open a new input file, save the delete commands as a text file, print a hard copy of the current plot, or exit from the program.

The Edit menu (Fig. 3) contains high-level functions used to edit the data. This menu provides the capability to cut (delete the currently selected) points, remove (delete) points that are off-scale for the current plot, select around periapsis (point of lowest altitude for planetary orbiters, such as Magellan), or undo (in reverse order) the previous delete operations.

The View menu (Fig. 4) allows the user to easily change the way that the data are displayed in the plot, primarily the plot scale. This menu provides the capability to refresh the plot, change the plot scale interactively, automatically select a scale to center the data in the plot, switch to the default scale, or perform a median absolute-deviation scaling operation.

The Options menu (Fig. 5) allows the user to change more details about what is displayed and how the program behaves. This menu provides the capability to change the plot label, change the plot symbols used for different stations, toggle the display of mean and sigma lines on the plot, change the behavior of the lasso, or change the trajectory file (PVFILE) for periapsis computation.

The Help menu (Fig. 6) provides additional on-line information for the user about how to use the program and its features. This menu provides additional text information on the purpose of the program, how the menus are structured and how they operate, and on the current version of XIDE.

C. Advanced GUI Features

In addition to the basic GUI capability provided by XIDE, there are several advanced GUI features that deserve special mention: the lasso feature, the plot-scale widget, and the magnifying glass.

The most frequently used advanced feature is the ability to lasso data points to be deleted with the mouse. Figure 13 shows an entire workstation screen with the XIDE window in the front. The residual plot is differenced (two-way minus three-way) Doppler data for the Magellan spacecraft from August 15, 1990 at 13:00:00 UTC to August 17, 1990 at 01:00:00 UTC. There are clearly some bad points in the lower half of the residual plot. These could be deleted one at a time by using the mouse to point at each and then clicking the left button. In fact, when this is done, the individual point changes to a plus symbol, and the color changes to red, which indicates that the point has been selected for deletion. At any time, the Cut menu item, or equivalently the middle mouse button, can be used to actually delete the points that have been selected; this causes them to disappear from the plot. Conversely, they can be returned to the plot by using the Undo menu item in the Edit menu.

A much more powerful concept for deleting a group of points is to select them with the lasso. By holding down the left mouse button and dragging the cursor across the plot, a curve of any reasonable shape can be drawn (for example, three quarters of a circle as shown in Fig. 13). When the left mouse button is released, an imaginary line is drawn between the two endpoints, and all the points inside the curve are selected (i.e., they change to a red plus symbol), and they can then be cut from the plot with the middle mouse button. The lasso is probably the single most effective feature of the XIDE program.

The second most important advanced feature is the Change Plot Scale dialog box (Fig. 14). This feature is implemented as a widget, which means that it is a GUI component that is reusable by other programs. This is not the only paradigm for changing the plot scale but it is an acceptably powerful one. As shown in Fig. 14, a miniature version of the residual plot is displayed. There are bars extending beyond the edge of the plot, two horizontal and two vertical. By using the mouse to touch one of these four bars with the cursor, the user can grab that bar by pressing the left mouse button, then drag the bar to the new desired location on the miniature plot, and then release the mouse button. As this is done, the appropriate numeric values are dynamically updated to reflect new minimum and maximum values of the x-axis date and time and the y-axis residual scale, which are displayed below the miniature plot. Also, the bars can only move in the appropriate horizontal or vertical direction and are bounded by the edge of the miniature plot. When the user has finished focusing on the data of interest, the Done button will return to the original plot (for example, as shown in Fig. 10); however, the new minimum and maximum values, which were selected without touching the keyboard, will be used for the residual plot.

The last advanced feature to be mentioned is the magnifying glass, shown in Fig. 15. In this case, the data set is two-way S/X-band Doppler data for the Ulysses spacecraft from January 9, 1991 at 00:00:00 UTC to April 25, 1991 at 00:00:00 UTC. As indicated by the label at the top of the plot in the XIDE window, there is a total of 4,402 after-the-fit points, and only those points used to construct this new orbit solution are displayed. When there are numerous such points, the fine structure of the data (for example, station dependency) can be quickly examined by pressing the right mouse button. A small square then appears on the plot where the cursor is, and a corresponding Zoom window with an expanded scale appears in the top-left corner of the screen. This window operates like a magnifying glass and follows the cursor as long as the right mouse button is pressed.

IV. Program Testing

The XIDE program was first used by the Magellan Navigation team, which possessed the prerequisite workstations. The team members must generate an estimated trajectory for Magellan on a daily basis, and they have been using XIDE operationally since the beginning of 1991; they use XIDE to edit radio metric data in the form of REGRES or PRESID files and produce delete commands for the ODP. The team members responsible for radio metric data editing estimate that by using XIDE, they have

reduced by 75–90 percent (from 4 hr to between 0.5 and 1.0 hr) the amount of time spent on these functions on a daily basis, as compared with the previous procedures used by the team on the UNISYS mainframe computer.

One of the important facets of testing the XIDE program was, in fact, demonstrating the portability of the source code. The XIDE program was developed from the beginning with this characteristic as one of the goals. Consequently, it was ported with relatively few changes from the original Sun UNIX environment to a Hewlett-Packard UNIX environment, and even to a DEC (Digital Equipment Corporation) VAX/VMS environment, while maintaining the same GUI from the user's viewpoint and essentially the same source code from the programmer's viewpoint.

The DSN Multimission Navigation team has been using the Pack Solve Analyze (PSA) program to perform DSN launch-support operations for several years [5]. This software is currently being used in a dedicated minicomputer environment based on the High Earth Orbiter (HEO) multimission navigation concept [6]. Because of the development and demonstrations of the XIDE prototype software in the workstation environment, the DSN Multi-

mission Navigation team has incorporated XIDE directly into the design of their second-generation interactive orbit-determination package.¹

V. Conclusion

The development of the XIDE program has proven to be successful from several perspectives. It was possible to integrate an existing navigation application with a GUI, specifically the OSF/Motif libraries, while still maintaining source-code portability. Once the prototype had matured to the point where it needed to be demonstrated, it was used experimentally in navigation operations by Magellan, an existing flight project, and ultimately was adopted as an operational tool. Finally, it was incorporated directly as an advanced capability into the design of the next generation of launch-support software for DSN multimission-navigation operations.

¹ N. A. Mottinger, J. Ellis, T. P. McElrath, P. R. Menon, K. E. Cridle, and B. Tucker, "Design and Functional Requirements for PSA with X-Windows, A Second-Generation Interactive Orbit Determination Package for HEO Launch Support," JPL Section 314 IOM 314.7-151 (internal document), Jet Propulsion Laboratory, Pasadena, California, September 24, 1991.

Acknowledgment

The author would like to thank Eric Slimko for providing the programming creativity to make the GUI for the XIDE program a reality while he was temporarily a member of the Navigation Software Development Group.

References

- [1] T. D. Moyer, *Mathematical Formulation of the Double-Precision Orbit Determination Program (DPODP)*, Technical Report 32-1527, Jet Propulsion Laboratory, Pasadena, California, pp. 1-160, May 15, 1971.
- [2] S. W. Thurman, "Application of Different Design Methodologies in Navigation Software and Development at JPL," *TDA Progress Report 42-101*, vol. January-March 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 39-49, May 15, 1990.
- [3] V. Quercia and T. O'Reilly, *X Window System User's Guide*, OSF/Motif Edition, Sebastopol, California: O'Reilly and Associates, Inc., 1991.
- [4] Open Software Foundation, *OSF/Motif Style Guide*, Revision 1.1, Englewood Cliffs, New Jersey: Prentice Hall, 1991.
- [5] V. N. Legerton and N. A. Mottinger, "PSA: A Program to Streamline Orbit Determination for Launch Support Operations," *TDA Progress Report 42-94*, vol. April-June 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 28-41, August 15, 1988.
- [6] J. Ellis, "HEO Multimission Navigation Concept," *TDA Progress Report 42-86*, vol. April-June 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 261-267, August 15, 1986.

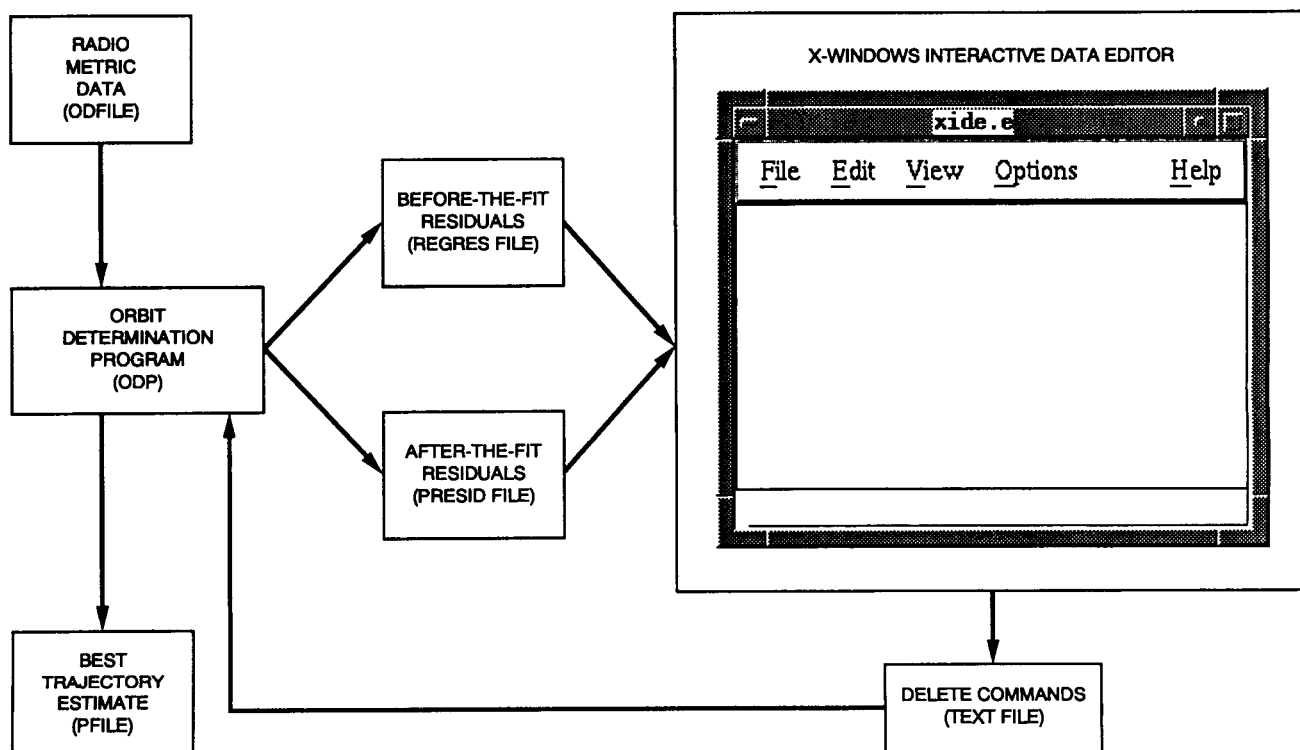


Fig. 1. X-Windows Interactive Data Editor overview.

<u>N</u> ew	Meta-N
<u>O</u> pen	Meta-O
Sa <u>v</u> e As	Meta-A
<u>P</u> rint	Meta-P
<u>E</u> xit	Meta-X

Fig. 2. File menu.

<u>R</u> efresh Plot	Meta-F
Ch <u>a</u> ng <u>e</u> <u>S</u> cale	Meta-S
C <u>e</u> nter <u>S</u> cale	Meta-C
<u>U</u> se Default Scale	Meta-D
Perfo <u>r</u> m <u>M</u> AD scale	

Fig. 4. View menu.

<u>C</u> ut	
<u>R</u> emove Offscale	Meta-R
S <u>e</u> lect Around <u>P</u> eriapsis	Meta-I
<u>U</u> ndo	Meta-U

Fig. 3. Edit menu.

Ch <u>a</u> ng <u>e</u> <u>L</u> abel	Meta-L
Ch <u>a</u> ng <u>e</u> <u>S</u> tation Symbols	Meta-T
<u>T</u> ogg <u>e</u> Lines	Meta-G
Ch <u>a</u> ng <u>e</u> <u>B</u> ehavior	Meta-B
Ch <u>a</u> ng <u>e</u> <u>P</u> V File	Meta-E

Fig. 5. Options menu.

On Purpose
On Menu
On Version

Fig. 6. Help menu.

New Meta-N
Open Meta-O
Save As Meta-A
Print Meta-P
Exit Meta-X

Fig. 7. File menu with some items grayed-out.

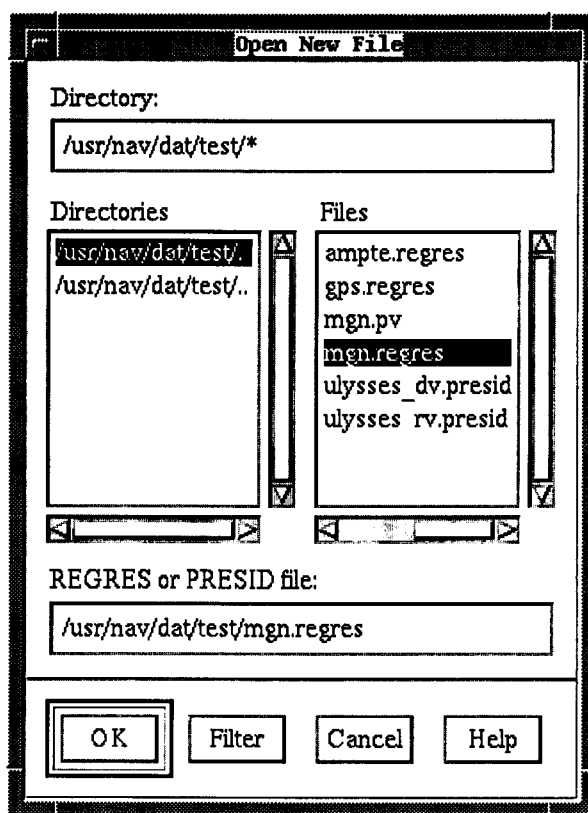


Fig. 8. Open New File dialog box.

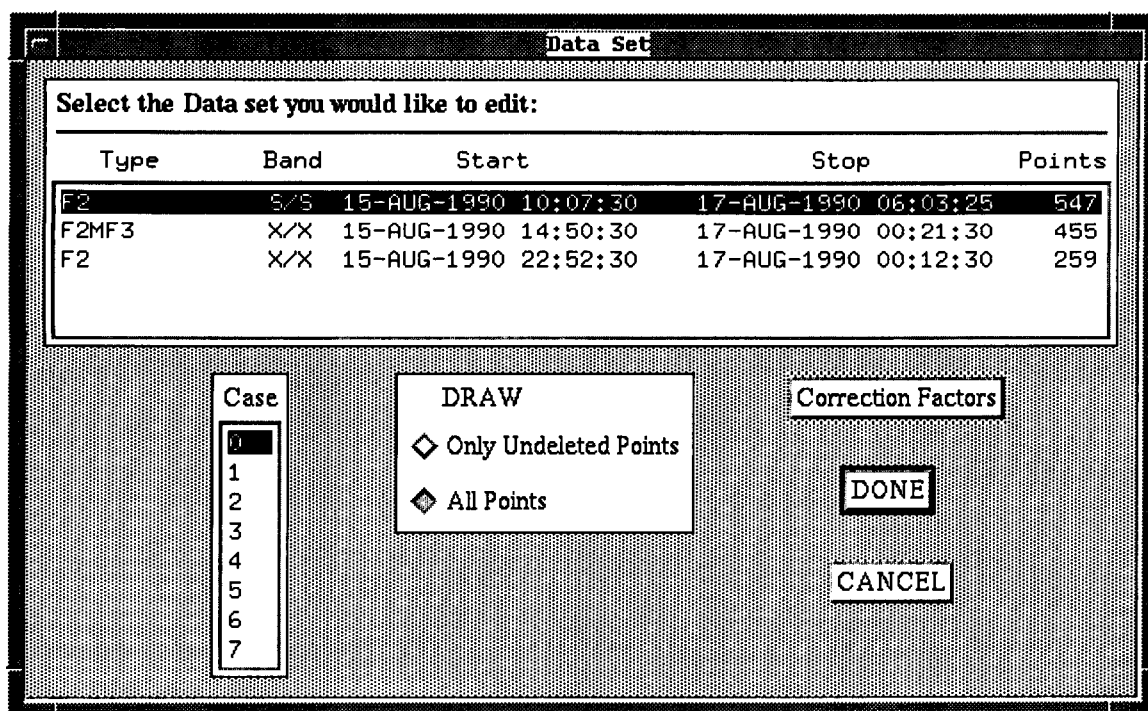


Fig. 9. Data Set dialog box.

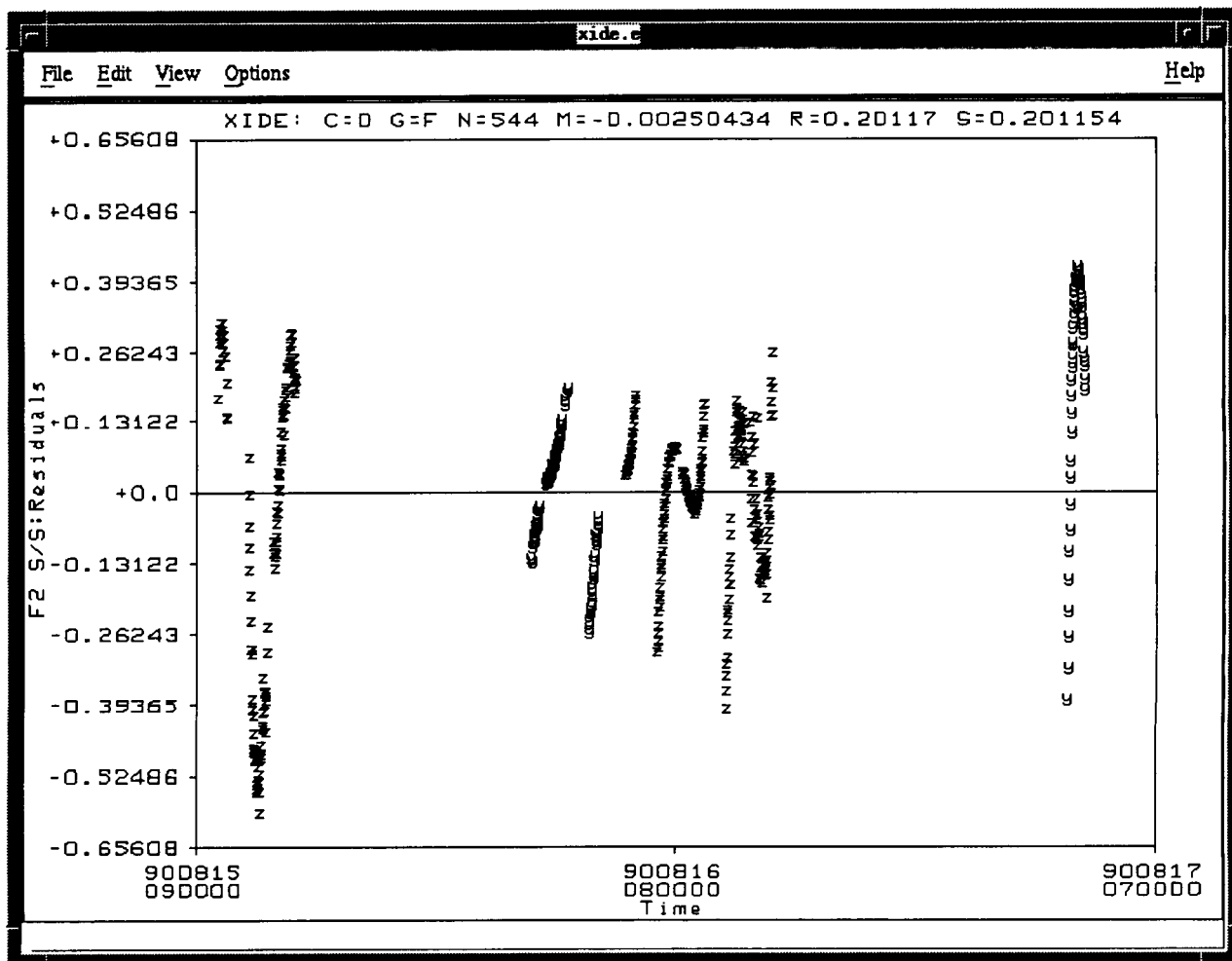


Fig. 10. Magellan Doppler residuals before the fit.

Write Delete Cards

Please enter the filename:

Format

☐ ODP Format

☐ ODE Format

Verb

☐ DELETE

☐ DELETE-ERASE

Station Limiter

☐ DSN

☐ None

Time Limiter

☐ FROM TO

☐ AT Only

Band Limiter

☐ UPLINK/DOWNLINK

☐ BAND

☐ None

Fig. 11. Write Delete Cards dialog box.

```

DELETE (F2MF3) BAND (X) FROM (90/08/15,14:50:29.9) TO (90/08/15,17:05:30.1)
      DSN (61) .
DELETE (F2MF3) BAND (X) FROM (90/08/15,21:22:29.9) TO (90/08/15,22:50:30.1)
      DSN (45) .
DELETE (F2MF3) BAND (X) AT (90/08/15,23:46:30.0) DSN (45) .
DELETE (F2MF3) BAND (X) AT (90/08/15,23:49:30.0) DSN (45) .
DELETE (F2MF3) BAND (X) AT (90/08/15,23:52:30.0) DSN (45) .
DELETE (F2MF3) BAND (X) FROM (90/08/15,23:54:29.9) TO (90/08/15,23:55:30.1)
      DSN (45) .
DELETE (F2MF3) BAND (X) FROM (90/08/16,00:25:29.9) TO (90/08/16,00:30:30.1)
      DSN (43) .
DELETE (F2MF3) BAND (X) AT (90/08/16,14:54:30.0) DSN (65) .
DELETE (F2MF3) BAND (X) AT (90/08/16,14:56:30.0) DSN (65) .
DELETE (F2MF3) BAND (X) AT (90/08/16,15:00:30.0) DSN (65) .

```

Fig. 12. Text file of delete commands.

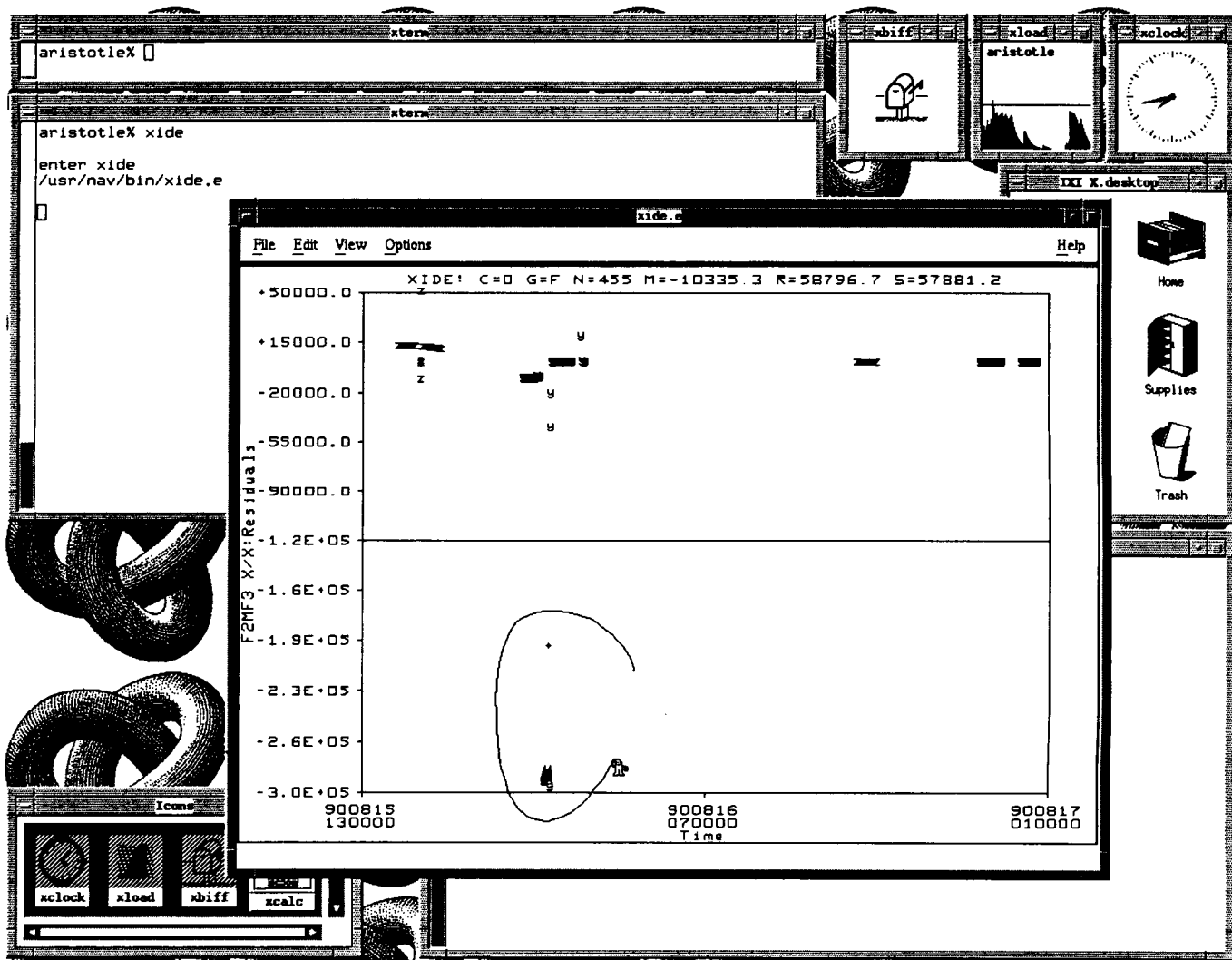


Fig. 13. Lasso feature in workstation environment.

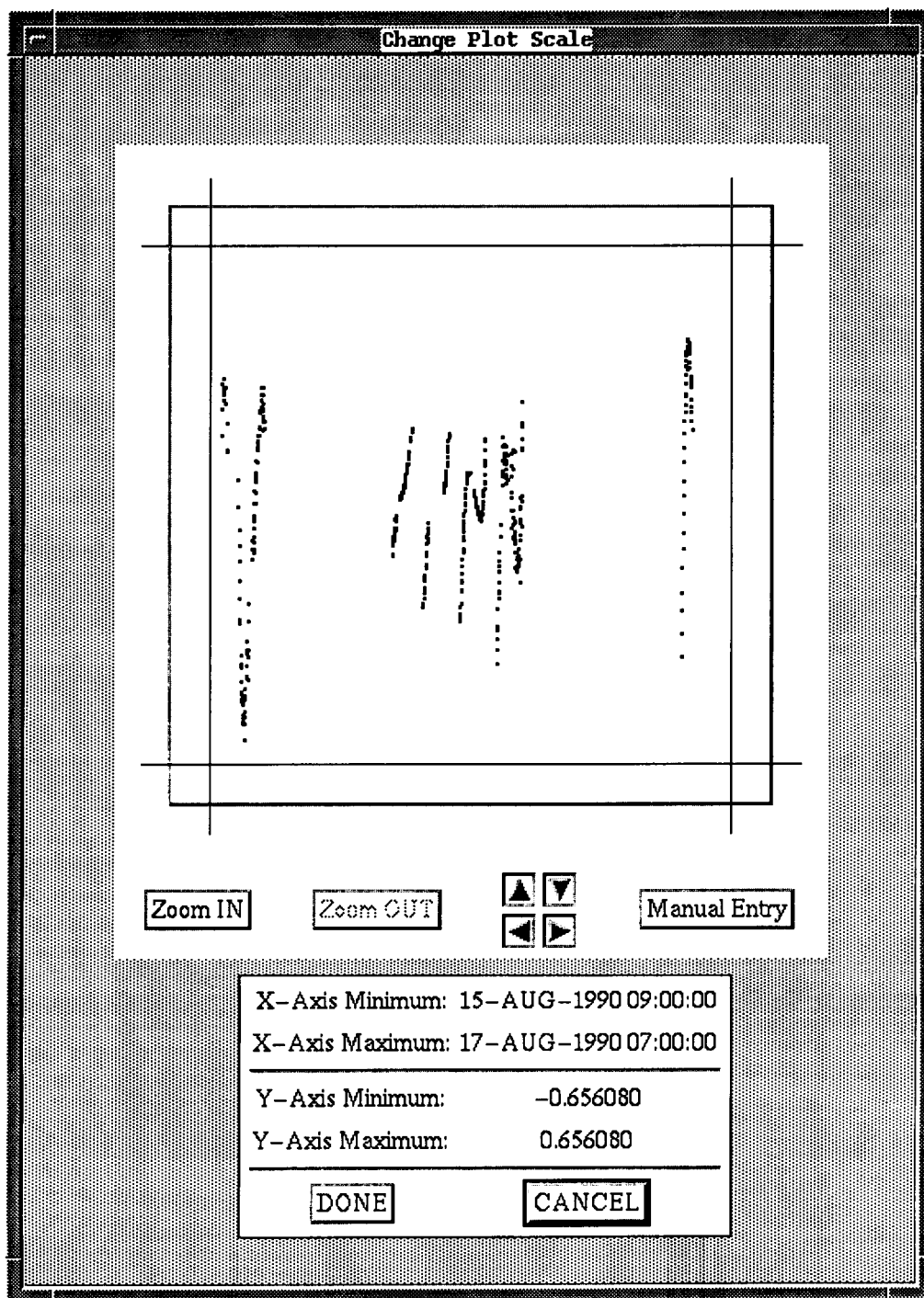


Fig. 14. Change Plot Scale dialog box.

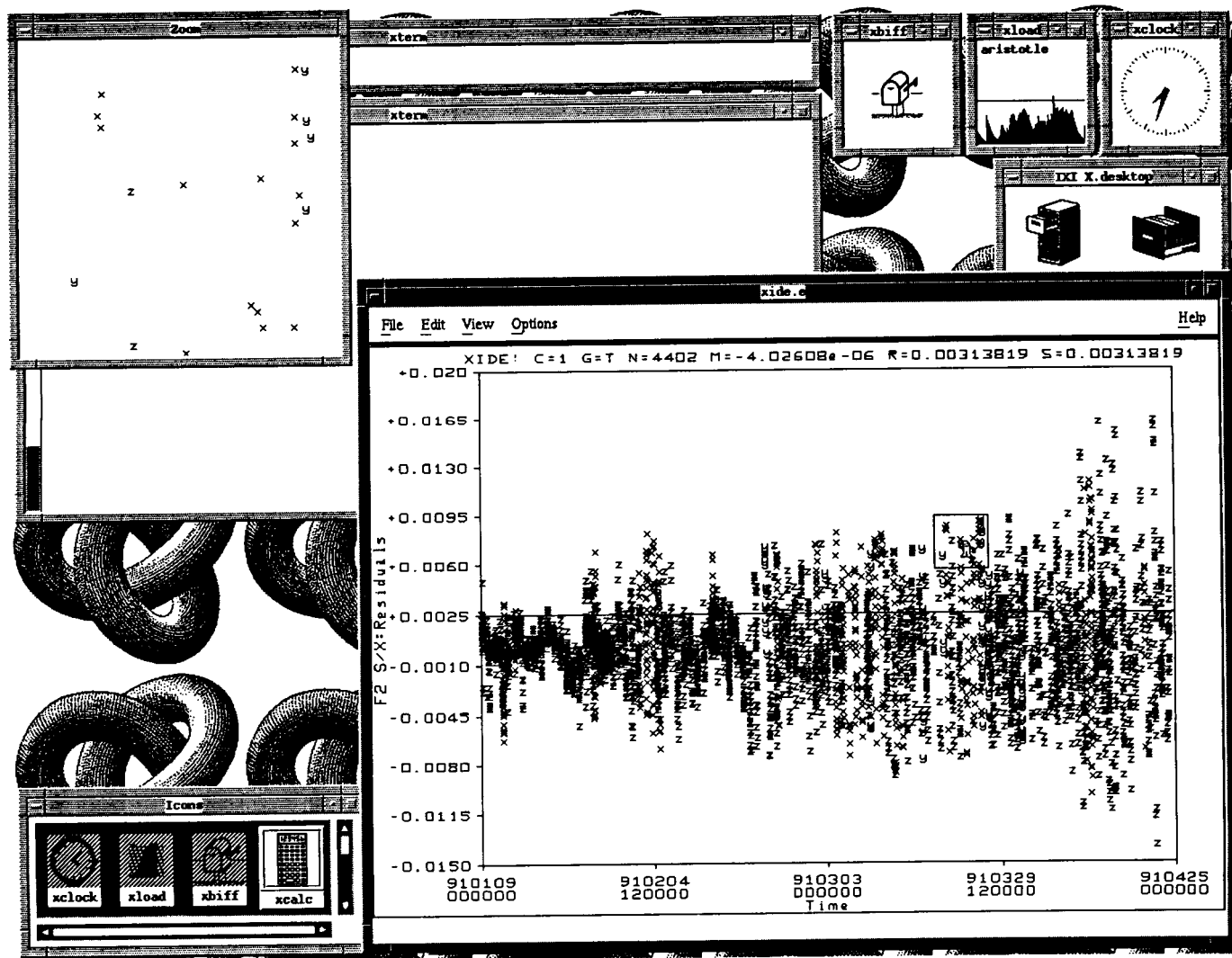


Fig. 15. Magnifying glass feature for Ulysses residuals.

57-59
p. 5
N92-24308
JJ574650

Alignment of the Atmospheric Visibility Monitoring Telescope

K. Cowles and D. Erickson
Communications Systems Research Section

Alignment of the first Atmospheric Visibility Monitoring telescope has revealed errors in mount design, mount manufacture, software, and electronics. This article discusses error sources and solutions, and realignment results. Alignment requirements for operation are also presented. The telescope now operates with the desired accuracy and repeatability.

I. Introduction

The Atmospheric Visibility Monitoring (AVM) telescope system requires fine alignment and repeatability to accurately move autonomously from one star position to another. The system measures starlight transmission through the atmosphere to provide joint visibility statistics for three sites in the southwestern United States. The sky will be scanned once an hour; the more stars that can be seen in an hour, the better the characterization of the atmosphere. In general, better alignment and repeatability decrease the time used to measure the illumination from each star. Initial alignment attempts revealed obvious problems with the telescope mount design and manufacture, and with the drive electronics. These problems prevented alignment. Specifically, the driver electronics caused the driver motors to run away randomly, declination bearing mounts had no means for adjustment during alignment, the declination wheel was continuously warping due to poor design, and declination and right ascension bearings were contaminated with dirt. Repairs to and replacements of each of these components allowed the mount to be aligned. The accuracy and repeatability of the mount's pointing was characterized.

During alignment, less obvious problems with the mount design and manufacture, and the driver software were revealed. Possible solutions, corrections made, and results are presented in this article.

II. Sources of Inaccuracy

Three problems inhibited accuracy and repeatability of the telescope pointing for the acquisition of stars: slippage during moves, preloading of the right ascension assembly due to the wheel not being concentric with the south-bearing shaft, and an improperly preloaded drive shaft for the declination wheel. Several measurements of error for travels of various distances were made to determine the effect of these problems on the ability to acquire stars.

A. Slippage

Slippage during moves from one star to another was caused by an abrupt start and stop of the frictional drive. In the original software, the stepper motor controllers received instructions to slew to a new position in the form of a direction and a constant speed. The sudden motion

from the stopped position to full speed and from full speed to stopped caused the wheels to slip on the drive shaft, introducing errors.

B. Right Ascension Drive Preloading

The right ascension wheel was balanced statically by measuring the torque required to move the wheel with a torque wrench. Magnetic counterweights were attached to the wheel in an attempt to compensate for torque differences. Some errors were not able to be eliminated using counterweights. These errors have been attributed to errors in the right-ascension wheel assembly, which was found not to be concentric with the South Pole bearing shaft. This caused uneven preloading, requiring more torque to turn the wheel at certain points. Figure 1 shows the torque necessary to move the right ascension wheel before it was repaired. The large peaks on the prerepair curve display the preloading effect. This error was also seen in alignment. When moving in right ascension where the wheel corresponded to a peak on the curve, larger positional errors were seen in locating the next star.

C. Declination Wheel Preloading

A static balance of the declination wheel was done separately from the alignment procedures. During the alignment procedures for the declination axis, the telescope appeared to be weighted more heavily to the south. Addition of weights to the north did not appreciably change the errors. Theoretically, if the telescope were heavier to the south, it would step too far south when moving south, and not far enough north when moving north. Changing the weights had little effect, however.

A compression spring preloads the drive shaft of the declination wheel. Decreasing the spring tension decreased the south-going error. If the spring tension was decreased to the point where there was still enough tension to drive the wheel without slippage, the telescope still pulled slightly to the south. By using counterweights and lowering the spring tension, the telescope could be moved between stars located between 20 deg N and 20 deg S with smaller errors. However, these errors were nonrepeatable, differing by up to 1.5 arcmin when traveling from north to south. This range limits the available area of the sky considerably. The problem has been attributed to improper preloading of the declination wheel, as seen in Figs. 2 and 3. Figure 2 shows the configuration of the declination drive. Figure 3 indicates the direction of force from the drive shaft onto the declination wheel. Figure 4 indicates the direction of the force that should be applied for proper preloading. Figure 5 indicates how the drive could be reconfigured to achieve this preloading.

III. Motivation for Improving Alignment

A telecompressor lens system is affixed in front of the camera to increase the field of view from 3.5×3.5 arcmin to 10.5×10.5 arcmin. A star can be 5.25 arcmin from the center of the viewing field and still be in the field of view of the initial frame, and 15.5 arcmin from center and still be found using a one-loop-square spiral search (assuming overlapping frames). Alignment of the telescope should provide for all stars to be found within 15 arcmin of the initial telescope move.

Better alignment allows more stars to be found on the first frame. If more stars can be found in the first frame, the number of stars that can be observed in an hour will increase. If a star cannot be found, the algorithm assumes a cloud is present, which would give incorrect data in the case of a star being farther than 15 arcmin from the initial telescope move. If the telescope system cannot find three stars in a row, it goes to the home position and reinitializes, a process which again takes time and introduces error in long travels. Improving the alignment of the telescope will result in a larger quantity of more reliable data.

IV. Implementation

A new version of the software ramps the speed of the motors between stopped and full speed, and vice versa, creating less wear on the motors and reducing errors. This should eliminate errors associated with sudden starts and stops in moves, thus improving repeatability in both the declination and right-ascension axes.

The right-ascension wheel assembly for the third telescope was remachined by the manufacturer to repair a damaged surface. While under repair, it was discovered that the south-pole bearing shaft was not concentric with the right-ascension wheel. It was determined that the preloading of the right-ascension wheel on the first telescope was probably caused by a similar problem. The first and second right-ascension wheels were then sent for repair. These wheels were also found not to be concentric. The center of the wheel for the first telescope was off by $3/8$ in. Repairs were made to both of the wheels.

The declination motor assembly and motor shaft have been redesigned and rebuilt, as indicated in Figs. 2-5, to provide proper preloading. A photograph of the present configuration is shown in Fig. 6.

V. Results

The curves in Fig. 1 showing conditions following repairs indicate that errors have been reduced by the ma-

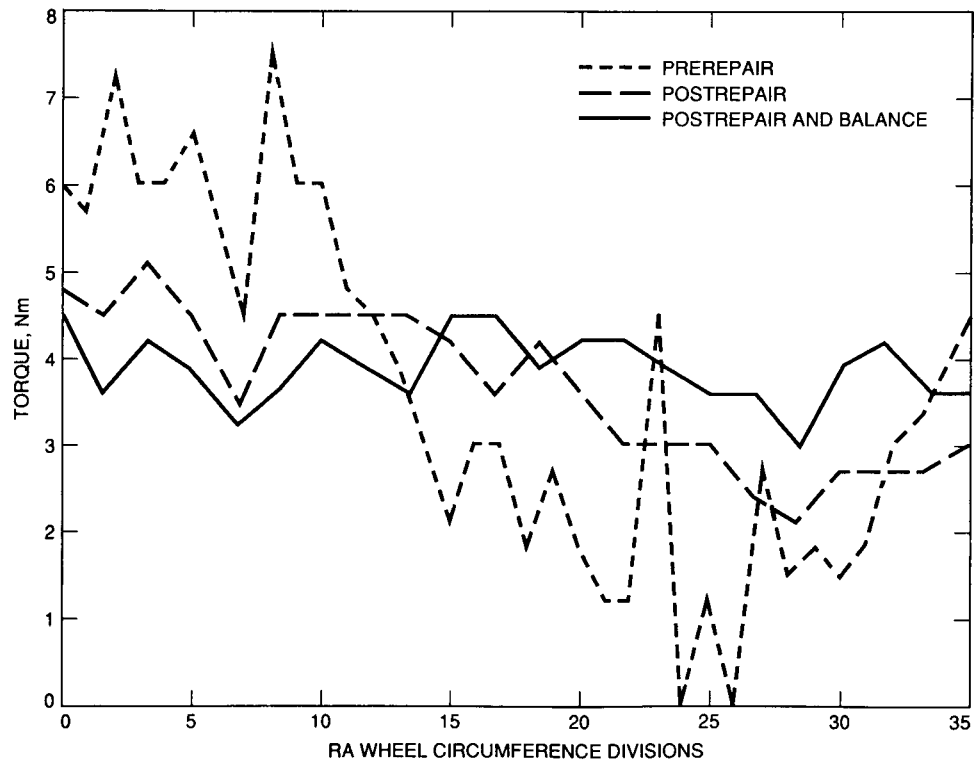


Fig. 1. The torque necessary to move the right-ascension wheel before repair, after repair, and after repair and balancing.

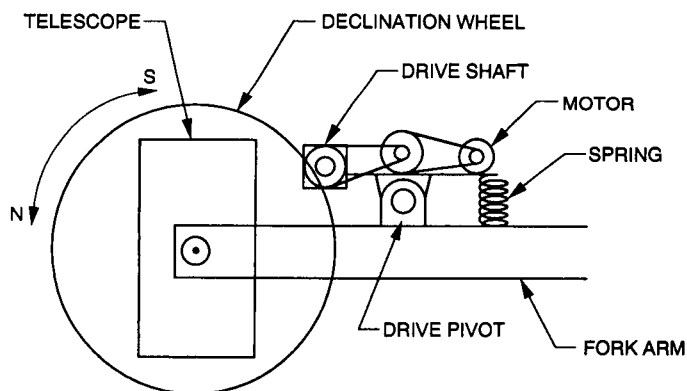


Fig. 2. Declination drive of the manufacturer's design.

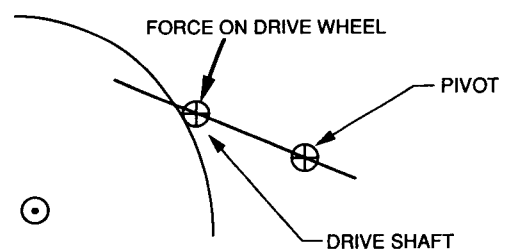


Fig. 3. Directions of force on the declination drive of the manufacturer's design.

chining of the right-ascension wheel. Balancing following the machining has resulted in a more even amount of torque necessary to drive the wheel.

Positional errors in moving between several bright stars were recorded after the right-ascension wheel modifications, software ramping modifications, and alignment. Fifty-five percent of the stars were found within the first frame of acquisition. Twenty-seven percent more stars were found with the use of a spiral search. However, 18 percent of the stars were not found because of large errors in the declination axis. Although star acquisition success was not measured in the same manner before these modifications were made, operator knowledge through repeated use of the system indicated that it was acting in a more accurate and repeatable manner.

Positional errors were measured again following modifications to the declination drive. Small and large moves between stars were performed to simulate typical telescope

operation. Longer movements did not indicate any increased difficulty in star acquisition. Sixty-nine percent of the stars were found on the first frame, and the remaining 31 percent were found within one spiral search.

VI. Conclusion

Flaws in design and manufacture have caused delays in integration and deployment of the first automated telescope mount system. The flaws were characterized during the initial alignment process. Repairs to the system have been made to eliminate these flaws. Tests were made to confirm the improvements to the system. The telescope now operates with the desired accuracy and repeatability, moving to stars within an error of 15 arcmin. Errors in telescope mount design and system alignment no longer inhibit star acquisition. Therefore, characterization of atmospheric transmission is not expected to include errors associated with equipment design and manufacture.

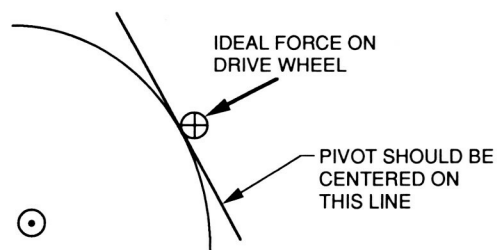


Fig. 4. Directions of force of the modified design for improved acquisition accuracy.

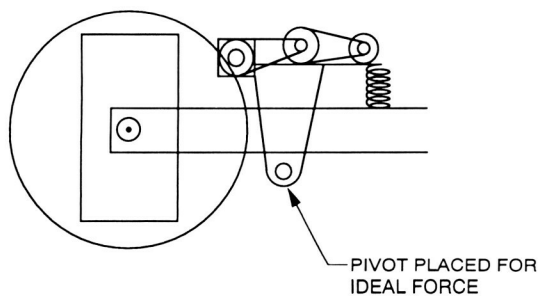


Fig. 5. Declination assembly configuration of the modified design for improved acquisition accuracy.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

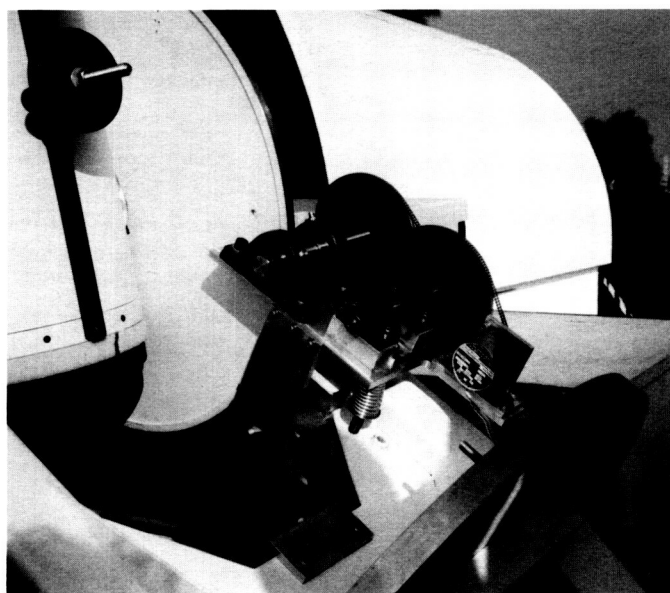


Fig. 6. The present configuration.

77-108
P-17 N92-24309

77-108
P-17 1450

Nonlinear, Nonbinary Cyclic Group Codes

G. Solomon¹

New cyclic group codes of length $2^m - 1$ over $(m - j)$ -bit symbols are introduced. These codes can be systematically encoded and decoded algebraically. The code rates are very close to Reed-Solomon (RS) codes and are much better than Bose-Chaudhuri-Hocquenghem (BCH) codes (a former alternative). The binary $(m - j)$ -tuples are identified with a subgroup of the binary m -tuples which represent the field $GF(2^m)$. Encoding is systematic and involves a two-stage procedure consisting of the usual linear feedback register (using the division or check polynomial) and a small table lookup. For low rates, a second shift-register encoding operation may be invoked. Decoding uses the RS error-correcting procedures for the m -tuple alphabet, i.e., the elements of the field $GF(2^m)$. Appendices A and B list $(m - j)$ -tuple codes for $m = 4, 5$, and 6 .

I. Cyclic Group Codes of Length $(2^m - 1)$ Over Binary $(m - 1)$ -tuples

Group codes of lengths up to 2^m over binary $(m - 1)$ -tuples are introduced and are shown to be cyclic and systematically encodable. These $(m - 1)$ -tuples are identified with an additive subgroup of the field $GF(2^m)$. These codes are not linear; that is, a codeword does not admit multiplication by a $GF(2^m)$ field element to yield another codeword.

Consider the field $GF(2^m)$ along with a primitive element β , which generates the $n = (2^m - 1)$ roots of unity. In addition, β is chosen with the following properties:

- (1) m odd: $\text{Tr } \beta^i = 0$ for $1 \leq i \leq m - 1$, where Tr denotes the linear field operator "trace." $\text{Tr } \beta = \beta + \beta^2 + \beta^4 + \dots + \beta^{2^{m-1}}$; thus, $\text{Tr } \beta \in GF(2)$, $\text{Tr } \beta^2 = \text{Tr } \beta$, and $\text{Tr } c\beta^2 = \text{Tr } \sqrt{c}\beta$, for $c, x \in GF(2^m)$.

- (2) m even: $\text{Tr } \beta^i = 0$ for $0 \leq i \leq m - 1$ except for a single odd integer p , $p < m$, and $\text{Tr } \beta^p = 1$.

The following are polynomials for β that satisfy the conditions (1) and (2) above for $3 \leq m \leq 12$.

m	Polynomial for β	Explanation
3	3 1 0	$(x^3 + x + 1)$
4	4 1 0	
5	5 3 0	
6	6 1 0	$(\text{Tr } \beta^5 = 1)$
7	7 3 0	
8	8 4 3 2 0	$(\text{Tr } \beta^5 = 1)$
9	9 5 0	
10	10 3 0	$(\text{Tr } \beta^7 = 1)$
11	11 9 0	
12	12 6 4 1 0	$(\text{Tr } \beta^{11} = 1)$

¹ Independent consultant to the Communications Systems Research Section.

Codes of length greater than 4096 are rarely invoked in present-day block-coding techniques. Do these properties extend beyond $m = 12$?

An element $c \in \text{GF}(2^m)$ may be represented by $c = \sum_{i=0}^{m-1} c_i \beta^i$. One may identify $\text{Tr } c$ by its binary representation (c_i) ; $0 \leq i \leq m-1$, and single out c_0 for m odd and c_p for m even. Thus, the binary value $\text{Tr } c$ is determined by only the trace-one position (0 or p) in its binary m -bit representation.

Choose an $(n, k; d)$ Reed-Solomon (RS) code over $\text{GF}(2^m)$ so that the codewords are values of sets of polynomials $P(x)$ with coefficients in $\text{GF}(2^m)$ of fixed highest degree $(n-d)$ or $(n-d-1)$. A codeword $\mathbf{a} = (a_j)$ is represented by the values of a polynomial $P_{\mathbf{a}}(x)$ so that $a_j = P_{\mathbf{a}}(\beta^j)$, $0 \leq j \leq n-1$.

Restrict $P_{\mathbf{a}}(x)$ for all codewords \mathbf{a} to an $(m-1)$ order subgroup of $\text{GF}(2^m)$ by stipulating that $\text{Tr } P(x) = 0$ for $x \in \text{GF}(2^m)$. Note that $P(x)$ as written here is generic for all $P_{\mathbf{a}}(x)$. The codes thus generated are cyclic group codes over $(m-1)$ -bit symbols and are systematically encodable for codes meeting the conditions in the main theorem.

Example 1. Take the RS code \mathbf{A} of dimension 5 over $\text{GF}(2^m)$, $\mathbf{a} \in \mathbf{A}$, $\mathbf{a} = (a_i)$, and $a_i = P_{\mathbf{a}}(\beta^i)$. The polynomials $P_{\mathbf{a}}(x)$ are of degree 4 with $\text{Tr } P_{\mathbf{a}}(x) = 0$, for all $x \in \text{GF}(2^m)$. For a general $P(x)$ and dropping the subscript \mathbf{a} , $P(x) = A + Bx + Cx^2 + Dx^3 + Ex^4$; $A, B, C, D, E \in \text{GF}(2^m)$. The condition $\text{Tr } P_{\mathbf{a}}(x) = 0$ gives $\text{Tr } A = 0$, $B^4 + C^2 + E = 0$, and $D = 0$. This code has the binary dimension $(m-1) + 2m$.

For $m = 3$, one gets binary dimension 8 or dimension 4 over 2-tuples; i.e., a $(7, 4; 3)$ code over binary doubles. This is a reduction from the $(7, 5; 3)$ RS code over binary triples!

There exists no integer dimension over $(m-1)$ -tuples for $m > 3$ since $(m-1) + 2m$ is not a multiple of $(m-1)$.

Example 2. Take an RS code of dimension 11 over $\text{GF}(16)$ but choose as the Mattson-Solomon (M-S) set the polynomials $P(x)$ of degree 11, and set the constant term equal to zero. $P(x) = \sum_{i=1}^{11} c_i x^i$; $\text{Tr } P(x) = 0$ leads to

$$\begin{aligned} c_1^8 + c_2^4 + c_4^2 + c_8 &= 0 \\ c_3^4 + c_6^2 + c_9 &= 0 \\ c_5 + c_5^4 + c_{10}^2 + c_{10}^8 &= 0 \\ c_{11}^2 + c_7 &= 0 \end{aligned}$$

The number of binary dimensions is $12 + 8 + 6 + 4 = 30$, which is dimension 10 over binary triples. Thus, the $(15, 11; 5)$ RS code over $\text{GF}(16)$ is transformed into the nonsystematic $(15, 10; 5)$ code over trace-zero elements of $\text{GF}(16)$.

Example 3. Similarly, the RS $(15, 7; 9)$ code over $\text{GF}(16)$ using polynomials of degree 6 from 0 to 6, using analogous techniques, gives the relations

$$\begin{aligned} \text{Tr } c_0 &= 0 \\ c_1^4 + c_2^2 + c_4 &= 0 \\ c_6 + c_3^2 &= 0 \\ c_5 &= 0 \end{aligned}$$

The binary dimension count is $3 + 8 + 4 = 15$, which yields a $(15, 5; 9)$ code over triples. Compare this with

- $(15, 5; 11)$ RS over 4-tuples,
- $(15, 5; 7)$ Bose-Chaudhuri-Hocquenghem (BCH) over $\text{GF}(8)$ and $\text{GF}(2)$, and
- $(15, 4; 10)$ BCH over $\text{GF}(4)$ (doubles).

These nonsystematic codes are cyclic. Examine the conditions of the coefficients as in Example 2; for example, when a codeword $P(x)$ is shifted r positions to the left, the new polynomial $P'(x)$ that describes this new cycled codeword becomes $P'(x) = P(\beta^r x)$. If $P(x) = \sum c_i x^i$, then $d_i = c_i \beta^{ir}$ and

$$\begin{aligned} d_1^8 + d_2^4 + d_4^2 + d_8 &= 0 \\ d_3^4 + d_6^2 + d_9 &= 0 \\ d_5 + d_5^4 + d_{10}^2 + d_{10}^8 &= 0 \\ d_{11}^2 + d_7 &= 0 \\ d_1^8 + d_2^4 + d_4^2 + d_8 &= \beta^{8r} (c_1^8 + c_2^4 + c_4^2 + c_8) \end{aligned}$$

A. General Construction of These Cyclic Codes

To construct integer dimension codes of lengths up to 2^m over $(m-1)$ -tuples with distance d , start with an RS code characterized by polynomials with M coefficients in $\text{GF}(2^m)$. These polynomials are to be of degree $n-d$ or degree $n-d+1$ (if the constant term = 0). The RS $(n, k; d)$ code has $M = k$, $n = 2^m - 1$, and $k = n - d + 1$.

B. Conjugate Classes

Consider the map $x \rightarrow x^2$ and define the sets M_i for the odd integers i . Make them distinct:

$$M_i = \{ i2^j \pmod{2^m - 1} : j = 0, 1, 2, \dots, m-1 \}$$

Each distinct set contains either m integers or some m_i integers where $m_i \mid m$. The M_i 's are ordered in increasing i . For $m = 4$, the sets are $M_1 = \{1, 2, 4, 8\}$; $M_3 = \{3, 6, 9, 12\}$; $M_5 = \{5, 10\}$; and $M_7 = \{7, 11, 13, 14\}$.

The condition $\text{Tr } P(x) = 0$ translates into a condition on the integers in M_i that occurs as powers of x in the M-S polynomial. For each distinct set M_i there is a linear dependency. If all M_i have the same number of elements m , the effective k -dimension narrows down to $\deg P - r$ where r is the number of independent M_i 's that occur in the powers of x . To yield an integer dimension, $m(\deg P - r)$ must be a multiple of $(m-1)$. Moreover, if $\deg P = n - d - 1$, one gets an additional dimension in this case since the condition $\text{Tr}(\text{constant}) = 0$ gives another set of $(m-1)$ -tuples.

Example 4. In this example, $m = 5$ and codes are over 4-tuples. Start with $(31, 27; 5)$ RS code over $\text{GF}(32)$. The M-S polynomials here are of degree 26 and include the constant term. $M_1 = \{1, 2, 4, 8, 16\}$; $M_3 = \{3, 6, 12, 24, 17\}$; $M_5 = \{5, 10, 20, 9, 18\}$; $M_7 = \{7, 14, (28), 25, 19\}$; $M_{11} = \{11, 22, 13, 26, 21\}$; and $M_{15} = \{15, (30), 29, 27, 23\}$. Integers in parentheses indicate powers omitted in the M-S polynomials.

With degree of $P(x) = 26$, code dimension $k = 27$, and $r = 6$. Here one has a $(31, 26; 5)$ code. This is obtained as follows: $26 - 6 = 20$ (5-tuples) and $20 \times 5 = 25 \times 4 = 25$ (4-tuples). Add the constant term to obtain 26 4-tuples. Since $\deg P(x) = 26$, the 4-tuple code obtained is the $(31, 26; 5)$ code. This code is cyclic and nonsystematic when encoded via the coefficients of the polynomials $P(x)$. There are ways to encode systematically.

Example 5. Using similar arguments, one obtains

$\deg P(x)$	m	r	Code over 4-tuples
22	23	6	$(31, 21; 9)$
18	19	6	$(31, 16; 13)$
13	13	5	$(31, 10; 19)$

The full list of length-31 codes appears in Appendix A.

C. Alternate Construction

Consider generator polynomial $g(x)$ with roots as powers of powers of β in either the set $\mathbf{A} = \{1, 2, 3, 4, \dots, 2t\}$ or $\mathbf{B} = \{0, 1, 2, 3, \dots, 2t-1\}$. Choose \mathbf{A} for m odd and \mathbf{B} for m even. Examine the conjugate classes M_i that intersect \mathbf{A} or \mathbf{B} . Let N_i be the number of elements in these M_i , and let A_i stand for the number in $M_i \cap \mathbf{A}$ or $M_i \cap \mathbf{B}$. There are $(N_i - A_i - 1)$ remaining independent constants for each i to be summed over the relevant i . Call this sum T : $T = \sum (N_i - A_i - 1)$. If $Tm = S(m-1)$ for S an integer, then the dimension of the entire $(m-1)$ -tuple t -error correcting code is $S + \dim$ of t -error correcting BCH code over m -tuples. \mathbf{A} or \mathbf{B} is chosen so that $Tm/(m-1)$ is an integer.

Example 6. Let $m = 4$ and β be a root of $x^4 + x + 1 = 0$. Let the generator polynomial contain $\mathbf{A} = \{\beta^i : 0 \leq i \leq 3\}$. $M_1 = \{1, 2, 4, 8\}$; $M_3 = \{3, 6, 9, 12\}$; $M_1 \cap \mathbf{A} = \{1, 2\}$; and $M_3 \cap \mathbf{A} = \{3\}$. $T = (4 - 2 - 1) + (4 - 1 - 1) = 3$ and $S = 3 \times 4/3 = 4$. One therefore has a contribution of four dimensions over 3-tuples. Add this to the BCH code of dimension 6, which does the regular job, and obtain a $(15, 10; 5)$ cyclic code over triples.

Example 7. Let $m = 7$; $\mathbf{A} = \{1, 2, 3, 4, 5, 6\}$; β is a root of $x^7 + x^3 + 1$. Start by asking for a 3-error correcting code.

$$M_1 \cap \mathbf{A} = \{1, 2, 4\}; \quad M_1 = \{1, 2, 4, 8, 16, 32, 64\}$$

$$M_3 \cap \mathbf{A} = \{3, 6\}; \quad M_3 = \{3, 6, 12, 24, 48, 96, 65\}$$

$$M_5 \cap \mathbf{A} = \{5\}; \quad M_5 = \{5, 10, 20, 40, 80, 160, 33, 66\}$$

Therefore, one can have $(3 \text{ of } 7) + (4 \text{ of } 7) + (5 \text{ of } 7)$ linearly independent binary dimensions $= 12 \times 7 = 14 \times 6$. Now consider the $(127, 106; 7)$ BCH code over 6-tuples, which contain none of the elements of M_i ; $i = 1, 2, 3$. Thus, one has a total of 120 dimensions over 6-tuples, or a $(127, 120; 7)$ code over binary 6-tuples. Compare this with the distance-7 RS code, which has dimension 121. Only one dimension has been lost in obtaining a cyclic group code over 6-tuples.

D. Decoding

Receive the codeword in 6-tuples and decode for three errors. The decoding algebra is performed in $\text{GF}(2^7)$ and the 6-tuples are read as trace-zero elements in their 7-tuple form.

E. Systematic Encoding

An example is presented here, followed by a theorem for general construction.

1. Construction of a Systematic Cyclic Group (15, 10; 5) Code Over Binary Triples. Let $g(x)$ be the generator polynomial of the (15, 11; 5) RS code over GF(16) with α a root of $x^4 + x + 1 = 0$.

$$g(x) = \prod_{i=1}^3 (x + \alpha^i) = x^4 + \alpha^{12}x^3 + \alpha^4x^2 + \alpha x + \alpha^6$$

Here the binary 4-tuple (a, b, c, d) represents the field element $u = a + b\alpha + c\alpha^2 + d\alpha^3$. Binary triples (a, b, c) are

∞	∞	∞	∞	∞	∞	∞	0	∞	∞	∞	6	8	7	4
∞	∞	∞	∞	∞	∞	∞	1	∞	∞	∞	7	9	8	5
∞	∞	∞	∞	∞	∞	∞	2	∞	∞	∞	8	10	9	6
∞	∞	∞	∞	∞	∞	∞	4	∞	∞	∞	10	12	11	8
∞	∞	∞	∞	∞	∞	∞	5	∞	∞	∞	11	13	12	9
∞	∞	∞	∞	∞	∞	∞	8	∞	∞	∞	14	1	0	12
∞	∞	∞	∞	∞	∞	∞	10	∞	∞	∞	1	3	2	14

This shows that the information set at the positions 0, 1, 2, 3, 4, 5, 6, 8, 9, and 10 for the triple code consisting of the elements $\{0, 1, 2, 4, 5, 8, 10, \infty\}$ is complete, since only ∞ at position 7 gives a codeword of the (15, 10; 5) code over triples.

2. Explanation. Note for an entry in the eighth position, labeled here (α^7) or "seventh" position (counting from 0 now), the trace-one elements occur in pairs in the parity positions. This is because the RS code has $(x + 1)$ as a factor in its generator polynomial $g(x)$, and so the sums of all symbols add to zero. Note too that all the possible seven ways an even number of trace-one elements can occur do occur. One may take any information set of 10 trace-zero symbols plus ∞ in the "seventh" symbol as a set of information symbols for the RS code generated by $g(x)$. One gets four parity symbols, which must contain an even number of trace-one elements.

Now add a unique codeword γ from the set **E** above to what was just generated. The codeword γ is chosen to eliminate the trace-one elements if they exist in the parity section to obtain a codeword with entries only having trace zero. Note that it is not necessary to store all seven words. In the set **E**, the triples 0, 1, and 8 in the "seventh"

treated as elements $a + b\alpha + c\alpha^2$ of GF(16). Note that in GF(16), $\text{Tr } \alpha = \text{Tr } 1 = 0$, so the triples belong to the subgroup with elements of trace zero.

Place the 10 information symbols in the first 11 entries, omitting the eighth position for the RS encoder. Now introduce a unique triple called γ in the eighth position and run all 11 triples through the RS encoder to generate four parity symbols, which are elements of trace zero in GF(16). What is this unique γ ?

The information triples correspond to the elements α^j , for $j = 0, 1, 2, 4, 5, 8, 10$, and ∞ , where α^∞ denotes the zero additive identity of the field. Note the following set **E** of seven codewords (in powers of α) of the RS code:

position give rise to words with trace one in position pairs (1, 3), (1, 2), and (1, 4). One gets all pair combinations by taking sums of these. Therefore, it is a basis. One need only store three words of size $(3 + 4 + 4 + 4 + 4) = 15$ bits and set up an algorithm for usage. This is a full systematic encoding since the information symbols are transparent and accessible. The general construction will be presented later along with the general case for $(m - j)$ -tuple codes of length $2^m - 1$, $j \geq 2$.

The cyclic group codes defined above can be encoded systematically. The encoding will be in two stages: one employing a linear feedback shift register (LFSR) encoder of the RS type and a ROM and perhaps even a second LFSR encoder. Start with an RS code of length $n = 2^m - 1$ over GF(2^m) with distance d . The M-S polynomials for the code are either $P(x) = \sum_{i=0}^{M-1} c_i x^i$ or $Q(x) = \sum_{i=1}^M c_i x^i$. A cyclic group code of dimension k over $(m-1)$ -tuples and distance $d = n - M + 1$ was defined by using the condition that $\text{Tr } P(x) = 0$ or $\text{Tr } Q(x) = 0$. Then M' free m -bit constants were obtained where $M'm = k(m-1)$ for $Q(x)$ and $(M' - 1)k = (m-1)(k-1)$ for $P(x)$. In order to systematically encode either case, a theorem is needed for finding a particular set of independent coordinates.

Theorem. Let \mathbf{A} be a $(2^m - 1, k; d)$ group cyclic code over binary $(m - 1)$ -tuples as defined above by a conditioned set on M' coefficients of M-S polynomials $Q(x) = \sum_{i=1}^M c_i x^i$, where $M'm = (m - 1)k$. If there exists a set of k coordinates with the property that the only codeword in \mathbf{A} that is zero at these positions is the all-zero codeword, then these coordinates can serve as information symbol positions; i.e., all codewords may be generated from the coordinate values there. Furthermore, if these m points are contained in the first M consecutive positions, one may fill in the $M - m$ missing positions with elements using a ROM and an algorithm and then encode all the n positions from the first M via the usual RS shift-register encoder. The proof is similar for the $P(x)$ type.

Proof: If one has such a set of k coordinates, each of the k^{m-1} possible values taken at these coordinates generates a distinct codeword of \mathbf{A} . Otherwise, there are at least two codewords of \mathbf{A} that correspond to some k -tuple. (Note that \mathbf{A} has k^{m-1} codewords.) If two such codewords are added together, one obtains a nonzero code-

word in \mathbf{A} with zeros at these m positions, which contradicts the hypothesis of the theorem. If the k positions are in consecutive M positions of the codewords, there may be ways to fill something in and use the available RS encoder technology. \square

A Caveat. If one starts from the polynomial construction of these codes, one has no guarantee that such a set of k coordinates exists. Consider a cyclic code generated by the M-S polynomials $P(x)$; $x = \beta^i$ for $0 \leq i \leq 14$. Then $P(x) = c_3 x^3 + c_3^2 x^6 + c_5 x^5$; $c_3 \in \text{GF}(16)$, $c_5 \in \text{GF}(4)$. This gives a code of dimension 2 over trace-zero elements (which are represented by triples). It is seen here that two coordinates do not define any triple codeword in a systematic manner.

3. Construction of the (15, 2; 12) Code. List the field elements in powers of β where β is a root of $x^4 + x + 1 = 0$ (∞ denotes the zero of the field). The triples are represented by $(1, 2, 4, 8, \infty, 0, 5, 10)$. Listed here are 35 of the 64 codewords corresponding to $P(\beta^i)$ for $0 \leq i \leq 14$

∞	2	4	1	8	∞	2	4	1	8	∞	2	4	1	8	(5 shifts)
∞	8	∞	8	0	10	2	5	1	∞	0	4	10	2	10	(15 shifts)
∞	∞	0	1	8	0	5	5	4	4	5	10	∞	2	1	(15 shifts)

Note that ∞ occurs in all pairs of positions and does not satisfy the hypothesis of the theorem. This code cannot be generated systematically from any two positions.

II. Cyclic Group Codes of Length $(2^m - 1)$ Over Binary $(m - j)$ -tuples, $j \geq 2$

Consider codes of length $2^m - 1$ over j -tuples, where $j \leq m - 2$ and $\gcd(2^j - 1, 2^m - 1) = 1$. The extension requires no new ideas or theorems but actual definitions and calculations. Appendix B lists a set of these new codes for $k = 5$ and 6. Appendices C and D are constructions of two different codes over binary triples.

To obtain new codes of length $2^m - 1$ over j -tuples, $j < m - 1$:

- (1) Represent the j -tuples as subgroups of the trace-zero elements of the field, with m odd and m even treated differently.
- (2) Characterize algebraically to preserve cyclicity.

(3) Systematically generate the codes by invoking the theorem.

(4) Note that if $j \mid m$, or $2^j - 1 \mid 2^m - 1$, these are no better than BCH codes. See the example in Appendix C.

Case 1 (m Odd). Find β primitive in the field $\text{GF}(2^m)$ such that $\text{Tr } \beta^i = 0$, for $1 \leq i \leq m - 1$. To obtain a subgroup of $(m - 2)$ -tuples, stipulate that for $\mathbf{a} \in \text{GF}(2^m)$, $\text{Tr } \mathbf{a} = \text{Tr } \mathbf{a}\beta^{-1} = 0$. To generate subgroups of order $(m - j)$, continue adding $\text{Tr } \mathbf{a}\beta^{-i+1} = 0$, $i \leq j$ to the previous conditions.

Case 2 (m Even). Find β primitive such that $\text{Tr } \beta^j = 0$, $0 \leq j < p$ with p as close to $m - 1$ as possible. To obtain a subgroup of $(m - 2)$ -tuples, stipulate that for $\mathbf{a} \in \text{GF}(2^m)$, $\text{Tr } \mathbf{a} = \text{Tr } \mathbf{a}\beta = 0$. To generate subgroups of order $(m - j)$, continue adding $\text{Tr } \mathbf{a}\beta^{i-1} = 0$, $i \leq j$ to the previous conditions.

A. Cyclicity

The above conditions ensure cyclicity. For m odd and $j = m - 2$, $\text{Tr } P(x) = \text{Tr } P(x)\beta^{-1} = 0$ for the appropri-

ately chosen β . For the polynomial $P(x) = \sum c_i x^i$, $i = 1, 2, 4$, and 8 , the above conditions lead to the equations

$$c_1^8 + c_2^4 + c_4^2 + c_8 = 0$$

$$\beta^{-8}c_1^8 + \beta^{-4}c_2^4 + \beta^{-2}c_4^2 + \beta^{-1}c_8 = 0$$

Shift the $P(x)$ codeword r positions to the left. The new set of coefficients for the shifted word is $P'(x) = \sum d_i x^i$, $d_i = c_i \beta^r$. A simple calculation verifies that the d_i 's meet the conditions above.

B. Systematic Encoding

The theorem states that a set of k coordinates can serve as a basis for the group code if there are no nonzero codewords over j -tuples that are zero at these coordinates. Let these k coordinates be the information coordinates. Now if these k 's happen to fall in any set of k' consecutive coordinates where k' is the degree of the M-S polynomial of the linear cyclic code over the full field, one may encode

$$\begin{array}{cccccccccccccccc} \infty & \infty & \infty & \dots & \infty & \infty & \infty & \infty & 0 & \infty & \infty & 22 & 25 & 18 & 23 \\ \infty & \infty & \infty & \dots & \infty & \infty & \infty & 0 & \infty & \infty & \infty & 13 & 27 & 22 & 28 \end{array}$$

From the above, it can be shown that if the coordinates 0-22, 25, and 26 (skipping 23 and 24) are chosen as information positions, they are a basis for group-code generation. To see this, multiply both sequences by the eligible triples and verify that one does not obtain a codeword that contains all triples.

2. Encode. Take any information sequence of triples, place ∞ in positions 23 and 24, and generate a codeword of the full RS code. Examine the four parity symbols for nontriple elements and add the proper unique codeword of the form

$$\infty \infty \infty \infty \infty \dots ab \infty \infty cdef$$

so that a and b are triples chosen such that c, d, e , and f have ones in same first two positions as the first codeword. That such a word exists is guaranteed below. First, count the dimension of the triple code.

3. Dimension Counting. The conjugate class of 1 2 is 1 2 4 8 16. Since three dimensions are needed to get one, only one dimension is extracted from this class. Now

systematically by using the encoding shift register plus the small size ROM template that fills in the missing $(k' - k)$ values plus the remaining parity symbols. The technique is similar to the $(m - 1)$ -tuple case.

1. Example of Construction of a (31, 25; 5) Code Over Binary Triples. Construct the field $GF(32)$ by choosing β as a root of $x^5 + x^3 + 1 = 0$. The trace-zero elements are 1, 2, 4, 8, 16, 3, 6, 12, 24, 17, 15, 30, 29, 27, 23, and ∞ (the zero identity of the field). The trace-one elements are 0, 5, 10, 20, 9, 18, 11, 22, 13, 26, 21, 7, 14, 28, 25, and 19. Take a set of elements $\mathbf{a} \in GF(32)$ with $\text{Tr } \mathbf{a} = \text{Tr } \mathbf{a}\beta^{-1} = 0$. There are eight elements here and they are (in powers of β) 2, 3, 4, 16, 17, 24, 30, and ∞ . Represent $GF(32)$ by 5-tuples (a, b, c, d, e) corresponding to $\mathbf{a} = a + b\beta + c\beta^2 + d\beta^3 + e\beta^4$. The triples to encode are now $(0, 0, c, d, e)$ in this representation. The generator polynomial for the (31, 27; 5) RS code over $GF(32)$ is given by $g(x) = \prod_{j=0}^3 (x + \beta^j)$. Explicitly, $g(x) = x^4 + \beta^{11}x^3 + \beta^{20}x^2 + \beta^{14}x + \beta^6$.

Consider the RS codewords given by

consider the class containing three dimensions. There are four elements left, so two can be extracted, giving a total of 3×5 bits, which in symbols equals five (triples) in dimension. There is a BCH code of dimension 20 that has $d = 5$, so the total triple dimension is $20 + 5 = 25$. To verify the encoding, first an even number of elements \mathbf{a} with $\text{Tr } \mathbf{a} = 1$ must occur in the four parity symbols. The RS encoder has $(x + 1)$ in the generator polynomial. Secondly, an even number of $\text{Tr } \mathbf{a}\beta^{-1} = 1$ must occur for the same reason. There are at most 8×8 such combinations. Now the elements a and b in these positions give rise to 64 different word possibilities and they must all yield different possible patterns, otherwise one would get a codeword of triples generated by $\infty \infty \infty \infty \infty \dots ab \infty \infty cdef$, which, as has been seen, is impossible by sheer calculation.

C. Optimality of These Codes

One wonders how efficient or optimal these codes are. One can compare them with RS codes having the same number of parity checks and see how they differ, or look at BCH codes that have the same number of parity checks. Then there is always the simple Hamming bound to fall back on. For a length $n = 2^k - 1$, and field elements of

$q = n + 1$ elements, the t -error correcting code will have dimension r where

$$q^r \leq q^n / \left[1 + nq + \binom{n}{2} q^2 + \cdots + \binom{n}{t} q^t \right]$$

where $\binom{n}{j} = n! / j!(n - j)!$, which approximates roughly to

$$q^r = q^n / \left[\binom{q}{t} q^t \right] = q^{n-t} / \binom{q}{t}$$

or $r \leq n - t - \log_q \binom{q}{t}$. It is known that for the RS codes, $r = n - 2t$. Now for j -tuples, $q = 2^j$ and $n = 2^m - 1$. An examination of the codes (15, 10; 5) over triples, (31, 26; 5) over 4-tuples, and (63, 56; 7) over 5-tuples shows that they are close to the Hamming bound.

Appendix A

Cyclic Codes of Length $2^m - 1$ Over $(m - 1)$ -tuples

Cyclic Group Codes of Length 15 Over Binary Triples:

$(15, 5; 9), (15, 10; 5), (15, 7; 7)$

Cyclic Group Codes of Length 31 Over Binary 4-tuples:

$(31, 6; 23), (31, 10; 19), (31, 16; 13), (31, 21; 9), (31, 26; 5)$

Cyclic Group Codes of Length 63 Over Binary 5-tuples:

$(63, 7; 51), (63, 12; 45), (63, 19; 37), (63, 24; 31), (63, 28; 27), (63, 33; 25),$
 $(63, 36; 21), (63, 42; 17), (63, 47; 13), (63, 51; 9), (63, 56; 7)$

Appendix B

Cyclic Codes of Length $2^m - 1$ over j -tuples, $j \leq m - 2$

$m = 5, j = 3$. Codes over binary triples:

(31, 6; 21), (31, 10; 15), (31, 16; 11), (31, 21; 7), (31, 25; 5)

$m = 5, j = 2$. Codes over binary doubles:

(31, 6; 15), (31, 11; 13), (31, 16; 9), (31, 20; 7)

$m = 6, j = 4$. Codes over binary 4-tuples:

(63, 7; 47), (63, 12; 39), (63, 18; 31), (63, 22; 27), (63, 28; 25), (63, 30; 21),
(63, 35; 21), (63, 41; 15), (63, 47; 11), (63, 51; 9), (63, 57; 5)

Appendix C

Construction of the (15, 7; 7) Group Cyclic Code Over Triples

Let β be a root of $x^4 + x + 1$, with β a primitive generator of GF(16). List the field elements as powers of β and let ∞ denote the additive 0 of the field. The trace-zero elements as integer powers of β are $\infty, 0, 1, 2, 4, 8, 5$, and 10.

Construct the M-S polynomials of degree 8, which give rise to the (15, 9; 7) RS code. Set the coefficient of x^7 equal to zero and consider the check polynomial

$f(x) = \prod (x + \beta^j); j = 0, 1, 2, 4, 8, 3, 6, 5$, and $f(x) = (x + 1)(x^4 + x + 1)(x + \beta^3)(x + \beta^6)(x + \beta^5)$. This polynomial written in ascending powers of x has coefficients [14, 0, 7, ∞ , 7, 14, 4, 4, 0].

The following codewords illustrate that the information symbol coordinates can be chosen as 0, 1, 2, 3, 4, 5, and 7. (See the theorem.)

∞	∞	∞	∞	∞	∞	0	∞	4	6	14	10	1	11	5
∞	∞	∞	∞	∞	∞	1	∞	5	7	0	11	2	12	6
∞	∞	∞	∞	∞	∞	2	∞	6	8	1	12	3	13	7
∞	∞	∞	∞	∞	∞	4	∞	8	10	3	14	5	0	9
∞	∞	∞	∞	∞	∞	5	∞	9	11	4	0	6	1	10
∞	∞	∞	∞	∞	∞	8	∞	12	14	7	3	9	4	13
∞	∞	∞	∞	∞	∞	10	∞	14	1	9	5	11	6	0

One sees that 0, 1, 2, 3, 4, 5, 7 can serve as an information set for the triples, since there does not exist a nonzero word that is zero at these positions. Still, an algorithm would be preferable so that the codewords can be systematically generated via the modified (15, 8; 7) RS code generated by $f(x)$. For this, a means is needed of generating a triple value at position 6 that depends on the values at positions 0, 1, 2, 3, 4, 5, and 7. This can be achieved by using a ROM indexed by position and triple value at that position. This would be a 9-bit ROM (3 bits for position, 3 bits for the triple at that position, and 3 bits for the triple to be placed in position six). To obtain this ROM requires calculation beforehand.

The contents are listed with positions labeled from 0 to 7 (omitting 6), and triples listed by values (a, b, c) corresponding to $a + b\beta + c\beta^2$. In Table C-1 are shown only 21 such listings at the 0, 1, and 2 powers of β since these

serve as a basis for the trace-zero elements. Thus, the 9-bit ROM generates for each triple at each information position the values that are added to form an element placed in the sixth position of the encoding shift register.

A second encoding technique arises from this calculation and is guaranteed by the theorem. Place the information triples along with ∞ in the sixth position and generate the remaining seven 4-tuples in GF(16). From the 4-tuples, extract a 7-tuple corresponding to the trace-one coefficient of β^3 . Place this 7-bit element in a 7×3 ROM, which generates a triple to cancel the trace-one elements. Have this triple index a second ROM, which generates seven triples (see the table, which verifies the conditions of the theorem) to be added to the seven-triple portion of parity that has been generated.

Table C-1. Listings at the 0, 1, and 2 powers of β .

Position	Value	Parity at 6
0	0	5
0	1	∞
0	2	4
1	0	0
1	1	4
1	2	0
2	0	5
2	1	0
2	2	5
3	0	0
3	1	4
3	2	0
4	0	5
4	1	0
4	2	5
5	0	10
5	1	4
5	2	1
7	0	∞
7	1	0
7	2	8

Appendix D

(15, 12; 3) Cyclic Code Over Binary Doubles

There are several ways to obtain this code. The standard method involves using the BCH code over $GF(4)$. The generator polynomial has roots $1, \beta$, and β^4 , where β is a root of $x^4 + x + 1 = 0$. An alternative is to identify the binary doubles with the set of elements in $GF(16)$ (a, b) corresponding to $a + b\beta$. The (15, 12; 3) code is then made up of the BCH (15, 10; 4) over $GF(4)$ with generator polynomial $x^5 + x^4 + x^2 + 1$.

Add one of the 15 codewords generated by $(x + \beta^7)(x + \beta^{11})(x + \beta^{13})$ using the conditions coming from the above exposition. These words turn out to be any cyclic shift of

$$1, \beta, 1, \beta, \beta^4, 1, \infty, \beta, 1, \infty, \infty, \beta, \beta^4, \beta^4, \beta^4$$

Note that these binary doubles correspond to $\infty, 1, \beta$ and β^4 . It would be interesting to see if these codes are isomorphic once the two sets of binary doubles are identified. The BCH code has a simpler encoding.

N92-24310

JJ574450

P-13

Real-Time Antenna Fault Diagnosis Experiments at DSS 13

J. Mellstrom and C. Pierson

Ground Antennas and Facilities Engineering Section

P. Smyth

Communications Systems Research Section

This article describes experimental results obtained when a previously described fault diagnosis system was run on-line in real time at the 34-m beam waveguide antenna at DSS 13. Experimental conditions and the quality of results are described. A neural network model and a maximum-likelihood Gaussian classifier are compared with and without a Markov component to model temporal context. At the rate of a state update every 6.4 seconds, over a period of roughly 1 hour, the neural-Markov system had zero errors (incorrect state estimates) while monitoring both faulty and normal operations. The overall results indicate that the neural-Markov combination is the most accurate model and has significant practical potential.

I. Introduction

In previous articles the problem of fault diagnosis of antenna pointing systems has been discussed in great detail [1,2]. Briefly, the problem is that of identifying whether an antenna pointing system is operating normally and, if not, which particular fault has occurred. The information available to the classifier system consists of time-series data from various sensor points in the antenna's pointing servomechanism. Identifying a fault condition is rendered nontrivial by the fact that feedback, redundancy, nonlinearities, and a considerable amount of noise from external disturbances are all present in the servo-control loop. This results in a masking of the underlying cause of a particular problem.

In [2] a hybrid signal-processing/neural-network architecture was applied to the problem. Specifically, various

characteristic features were extracted from 4-second blocks of the multichannel data. The most useful features for discriminatory purposes were found to be autoregressive coefficients and standard deviations (estimated for certain channels). In [2] it was further shown that classification accuracies of about 90 percent (on test data sets independent of the training data) were achievable using a neural network model, where there were four classes to be predicted: normal, tachometer failure, noisy tachometer, and compensation loss. By taking into account time correlations in the data, accuracies of about 98 percent were attainable.

The results in [2] were obtained using off-line postprocessing of data collected from the elevation axis drives of the 34-m beam waveguide (BWG) antenna at DSS 13. Specifically, the time series sensor data were recorded directly in digital format by the LabView data acquisition

software running on a Macintosh II computer. At a sampling rate of 50 Hz with 12 channels of interest, this resulted in large volumes of data being recorded; hence, for practical reasons, only 5 minutes worth of data were recorded for each class on any given day of data acquisition, since even a 5-minute batch of such data typically requires about 0.4 Mbyte of storage memory. As described in [2], the 5-minute segments of data were then downloaded to a Sun workstation computer to be used for training and testing various fault diagnosis algorithms. This *off-line* analysis led to the development of specific pattern recognition algorithms described in [2].

In this article, an *on-line* experiment is described where a particular classification model was implemented as part of the data acquisition software, allowing direct on-line estimation and classification at the antenna itself. The primary goal of this experiment was to determine if the pattern recognition system described in [2] could be implemented on-line at low cost and provide reliable estimates of normal and fault conditions in real time. The article describes the experiment setup, the results obtained, and concludes that the system has significant practical potential for antenna monitoring.

II. Implementation of the On-Line Fault Diagnosis Software

In [2] a fault diagnosis system was described which estimates various features from multichannel time-series data over consecutive windows, calculates a posterior probability estimate using these features, and then updates its estimate of the state of the system as a function of the current class estimates based on this window and the previous estimates. In [2] a number of different models were evaluated, having roughly comparable performance in terms of classification accuracy. For this experiment the authors chose to implement one particular model using the LabView software package. The LabView software package is intended for real-time data acquisition and analysis. Its primary advantage is that the system can be easily programmed and modified using an intuitive graphical user interface. Basic data acquisition capabilities, filtering, and signal analysis can be configured directly by editing a block diagram display of the signal flow in the system. For the purposes of this experiment the primary functions to be implemented included analog to digital conversion and buffering of the multichannel sensor data, estimation of the classification features on a block-by-block basis, implementation of the classification equations which estimate class probabilities as a function of the features, and a memory-based algorithm which updates a context-based estimate of the system state given each block.

A. Data Acquisition and Timing

The LabView software was set up to acquire k consecutive samples on each channel of interest where the data are sampled at a rate of f_s . For the experiments described in this article $f_s = 54$ Hz, which is sufficient given that all signals of direct interest are below 10 Hz. A block size of $k = 200$ was chosen providing a reasonable trade-off between being large enough to get reliable parameter estimates but not having too large of a delay between classification decisions. Hence although the time series data sampling rate is 54 Hz, the classifier is operating at a much slower rate since a decision every few seconds is deemed quite adequate for this type of application. For simplicity of implementation the data were not pipelined in the LabView software, which meant that data were sampled and collected for the period of τ_1 seconds (where $\tau_1 = k/f_s$) and then processed (parameter estimation, classification, and state update) for another τ_2 seconds. After this total interval of $\tau_t = \tau_1 + \tau_2$ seconds, the next block of sample data was acquired. Hence, the state estimates were updated every τ_t seconds. In the actual experiments to be described later, τ_2 was on the order of 2.7 seconds leading to a state update every 6.4 seconds. The "gaps" in the data of 2.7 seconds (between the 3.7-second blocks) were deemed inconsequential to the overall operation of the system, since no portion of the model relies on an assumption that the window blocks are contiguous in time. However, it should be noted that if this system were to be implemented in an operational mode, a pipelined scheme with no data gaps would be a more elegant solution.

B. Feature Estimation

The particular classification model chosen for the experiment uses 12 features. The first eight features are autoregressive-exogenous (ARX) model coefficients estimated from a model which has the antenna servo-controller elevation rate command as the exogenous (forcing) model input and the motor current of one of the elevation drive motors as the model output (see [2] for more details). These coefficients were estimated from the 200-sample blocks using a standard least-squares estimation algorithm. The other four features used as input to the classifier are: estimated standard deviation of both elevation drive motor currents, estimated standard deviation of the average tachometer sensor, and the estimated standard deviation of the difference of the two tachometer sensors.

C. Classification Using a Neural Network Model

The particular classification model implemented was a multilayer neural network model: a feed-forward network

with a single hidden layer of nonlinear "hidden" units. The model had 12 inputs (one for each of the features) plus a constant bias input set to 1. The particular model used in the experiments had 20 hidden units. The output layer consisted of four units, one for each class: normal, tachometer failure, tachometer noise, and compensation loss. An output *activation* vector $\mathbf{y}(t)$ at time t is calculated as a function of the input feature vector $\mathbf{x}(t)$ as described in detail in Appendix A of [2]. In particular this model can be implemented via two matrix multiplications and a componentwise nonlinear vector transformation, i.e.,

$$\mathbf{y}(t) = \gamma \left(\mathbf{W}_2 \cdot \gamma(\mathbf{W}_1 \cdot \mathbf{x}(t)) \right)$$

where \mathbf{W}_1 and \mathbf{W}_2 are the weight matrices between the input and hidden layers and hidden to output layers, and $\gamma(\mathbf{u}) = (\gamma(u_1), \dots, \gamma(u_n))$ where $\gamma(x) = 1/(1+e^{-x})$. The i th component of the output vector $\mathbf{y}(t)$ is interpreted as a rough estimate of the posterior probability of class i given the input feature vector $\mathbf{x}(t)$.

D. State Estimation With a Hidden Markov Model

The class probability estimates as produced by the classifier at each time t do not take into account the fact that faults are typically correlated over time. In [2] a heuristic scheme was described to reflect this prior information, where the current estimate of the system state was a function of both the present probability estimates and past probability estimates over a specified number m of past windows. In Appendix A a formal model of this time dependence is introduced using a Hidden Markov Model (HMM). The HMM approach has been successfully used in applications such as speech recognition to model temporal context. The n states of the system (in this case the normal and fault classes, $n = 4$) comprise the Markov model. For this application components of the Markov transition matrix \mathbf{A} (of dimension $n \times n$) are estimated *subjectively* rather than estimated from the data, since there is no reliable database of fault-transition information available at the component level. The *hidden* component of the model arises from the fact that one cannot observe the states directly, but only indirectly via a stochastic mapping from states to symptoms (the features). In speech recognition it is often desired to calculate the most likely sequence of states (perhaps phonemes) given some acoustic evidence. For a fault diagnosis application such as this it is sufficient to calculate at time t the probability that the system is in any particular state at that time. The state vector is denoted by $\mathbf{s}(t)$. As described in Appendix A the probability estimate of state i at time t can be calculated recursively via

$$\hat{\mathbf{u}}(t) = \mathbf{A} \cdot \mathbf{p}(\mathbf{s}(t-1))$$

and

$$p(s_i(t)) = \frac{\hat{u}_i(t)y_i(t)}{\sum_{j=1}^4 \hat{u}_j(t)y_j(t)}$$

where the estimates are initialized by a prior probability vector $\mathbf{p}(\mathbf{s}(0))$, the $s_i(t)$ are the components of $\mathbf{s}(t)$, $1 \leq i \leq n$, and the $y_i(t)$ are the outputs of the classifier as described in Section II.C.

E. Software Implementation

Implementation of the ARX parameter estimation, classifier equations, and HMM model estimation was straightforward using predefined functions in LabView. All of the above components of the algorithm were implemented, integrated with data acquisition software, and tested in the laboratory in about 3 working days.

As mentioned earlier, the time taken to process a single block of 3.7 seconds worth of data was estimated at 2.7 seconds. Of this time, it was estimated that 2.3 seconds were spent calculating the eight ARX coefficients and the standard deviation terms on the 200 samples, and 0.12 second were spent updating the screen displays; the actual classification equations (the neural network model and Markov equations) took only 0.08 second, and another 0.20 second went to other miscellaneous bookkeeping activities. It is interesting to note that the classification process itself is quite fast taking only 2.9 percent of the processing time, while the feature extraction takes up 85 percent of the time. These numbers are significantly slower than speeds which would be attainable for operational purposes were the system to be implemented using special-purpose signal processing hardware.

III. Experimental Conditions

The experiments were carried out in early November 1991 at the 34-m BWG antenna at DSS 13. Wind conditions were favourable for the duration of the experiments, i.e., no high wind conditions were encountered. Under each desired state (normal and fault classes) the antenna was driven at a typical DSN tracking rate (between 1 and 3 mil/degrees per second) along the elevation axis. Normal conditions corresponded to normal antenna operation. Fault conditions were simulated as described previously in [2], i.e., hardware faults were switched into the feedback loop of the elevation servo-control loop in a controlled manner. For safety reasons these faults were introduced when

the antenna was stopped rather than while it was moving (this precluded testing the response of the models in terms of transitions between faults). Each fault condition was then monitored after the antenna had been restarted. Hence, the experimental results in the next section reflect the fact that the data were collected in different batches in this manner. Faults were typically monitored for durations of 10 to 15 minutes and the total duration of data to be reported in the next section amounted to a total of about 1 hour of antenna monitoring. Longer duration monitoring was not practical due to various constraints such as other demands on the antenna and the amount of setup time required to get the monitoring system in place. Due to hardware problems it was not possible to simulate one of the faults described in [2], namely the tachometer noise condition. Hence, the results pertain to normal conditions, the tachometer failure, and compensation loss faults.

IV. Experimental Results

For purposes of comparison, the model implemented at DSS 13 is compared with a simple maximum-likelihood Gaussian classifier whose operation is described in Appendix B. The results with the Gaussian classifier were generated off-line using the recorded feature data. In addition, data were recorded with and without the Markov portion of the model to evaluate the effectiveness of this component. Hence, there are four types of models being compared: Gaussian, Gaussian followed by Markov, neural network, and neural network followed by Markov. Again it is emphasized that the results to be described pertaining to the latter two models were generated on-line in real time.

Table 1 summarizes the overall classification performance of each of the models on five different runs: two for normal conditions, two for the compensation loss fault, and one run for the tachometer failure fault. The units in the table correspond to windows of 3.7 seconds worth of sensor data spaced at 6.4-second intervals. The bottom row of the table tabulates the number of windows per run: hence, for example, 596 windows in total were analyzed, corresponding to about 1 hour and 3.6 minutes total duration. The numbers in the other four rows of the table (one row per model) indicate the number of windows misclassified by each model for each run, except for the final two columns, which give the overall number of misclassifications per model and the percentage misclassified, respectively. Clearly, from the final column, the neural-Markov model (implemented in real time) is the best model in the sense that no windows at all were misclassified. It is significantly better than the Gaussian classifier, which performed

particularly poorly under fault conditions. However, under normal conditions it was quite accurate, having only one false alarm during the roughly 30 minutes of time devoted to monitoring normal conditions. The effect of the Markov model is clearly seen to have beneficial effects, in particular reducing the effects of isolated random errors. However, for the compensation loss on day 316, the Markov model actually worsened the already poor Gaussian model results, which is to be expected if the non-Markov component is doing particularly poorly as in this case.

Figures 1 through 5 plot the estimated probability of the true class as a function of time for various models to allow a more detailed interpretation of the results. Note that, given that the true class is labelled i , the estimated probability of class i from the neural network corresponds to the *normalized* output of output unit i of the network at time t , i.e.,

$$\hat{p}_i(t) = \frac{y_i(t)}{\sum_{j=1}^4 y_j(t)}$$

while the Markov probabilities correspond to the estimates of state i , $p(s_i(t))$, as described earlier. Figure 1 corresponds to normal conditions on day 311 and compares the neural model with and without the Markov processing. Figure 1 demonstrates that the instantaneous probability estimates from the neural model have a large variation over time and are quite noisy. This is essentially due to the variation in the sensor data from one window to the next, since, as might be expected, signals such as motor current are quite noisy. In addition, a large glitch is visible at around 460 seconds. The neural model gives a low probability that the condition is normal for that particular window (in fact a large glitch such as this looks like a tachometer failure problem); however, the Markov model remains relatively unaffected by this single error. Overall, the stability of the Markov model is clearly reflected in this plot and should be advantageous in an operational environment in terms of keeping the false alarm rate to a minimum. Figure 2 is the same as Fig. 1 except that the data were obtained on day 316. Note that in both Figs. 1 and 2, at any particular instant the classifier assigns to the true class a probability of up to only 0.8 or 0.9. In contrast, by modelling the temporal context, the Markov model assigns a much greater degree of certainty to the true class.

Figures 3 and 4 compare the performance of the Gaussian and neural models on detecting the compensation loss fault. In Fig. 3, the variation in the Gaussian estimates

is quite marked. The Gaussian-Markov model combination, after some initial uncertainty for the first 90 seconds, settles down to yield reasonable estimates. However, the overall superiority of the neural-Markov model is evident. In Fig. 4 (the same fault on a different day), the neural and Gaussian models are compared directly. The variation in the Gaussian estimates is even worse. In fact, the resulting Gaussian-Markov estimates are almost random in nature and have been omitted for clarity.

Figure 5 also compares the Gaussian and neural models but on the tachometer failure fault. Once again the variation over time of the Gaussian estimates is unacceptably large, and again, the resulting Gaussian-Markov estimates have not been plotted due to their almost random nature, i.e., the variation over time on the estimates is so great that the Markov model cannot find any significant correlation.

In Fig. 6, the same information as that for Fig. 3 is plotted but in a different manner. Let \hat{p} be the probability estimate of the true class. Then $1 - \hat{p}$ is the effective error in the estimate; Fig. 6 is a logarithmic plot of this error term. Here, the overall dominance of the neural-Markov model is once again evident, as is the large variation in the Gaussian model.

V. Conclusion

The DSS 13 field results discussed in this article provide convincing evidence for the potential of a pattern recognition system to reliably monitor a DSN antenna drive system in real time. In addition, it should be noted that the monitoring system is unobtrusive and relatively easy to install due its use of inexpensive "off-the-shelf" PC-based hardware and software components.

Table 1. Summary of classification results for various models and runs during days 311 and 316 at DSS 13.

Model	Number of blocks misclassified					Overall misclassification statistics	
	Day 311		Day 316				
	Normal	Compensation loss	Normal	Tachometer failure	Compensation loss	Total no. of misclassifications	Total percent of misclassified
Gaussian	1	15	0	35	50	101	16.94
Gaussian+Markov	1	8	0	3	74	86	14.42
Neural	1	0	4	0	0	5	0.84
Neural+Markov	0	0	0	0	0	0	0.00
Total number of windows	128	60	152	126	130	596	

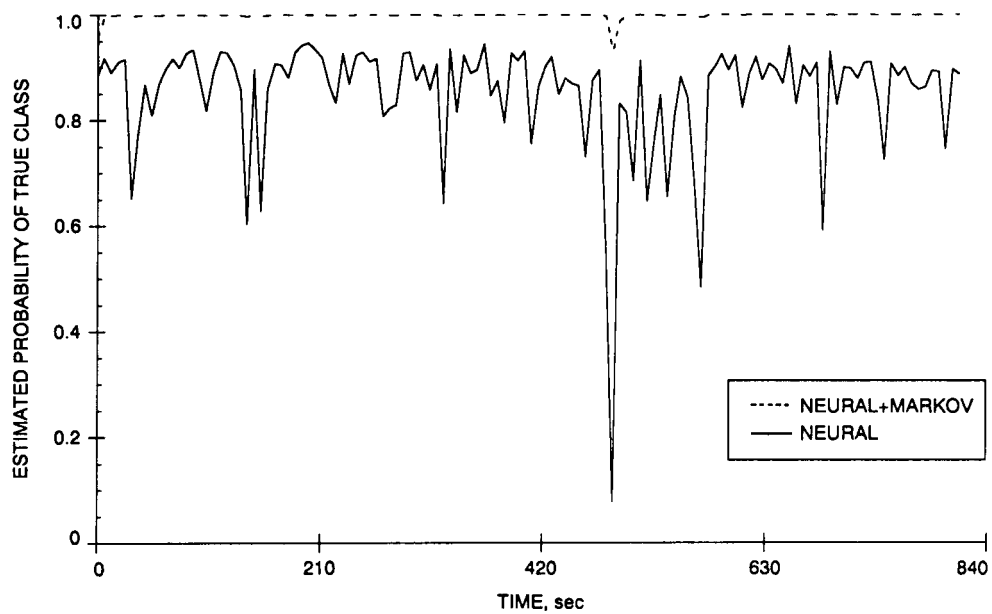


Fig. 1. Comparison of Markov and non-Markov models: estimate of probability of true class (normal conditions) as a function of time for day 311.

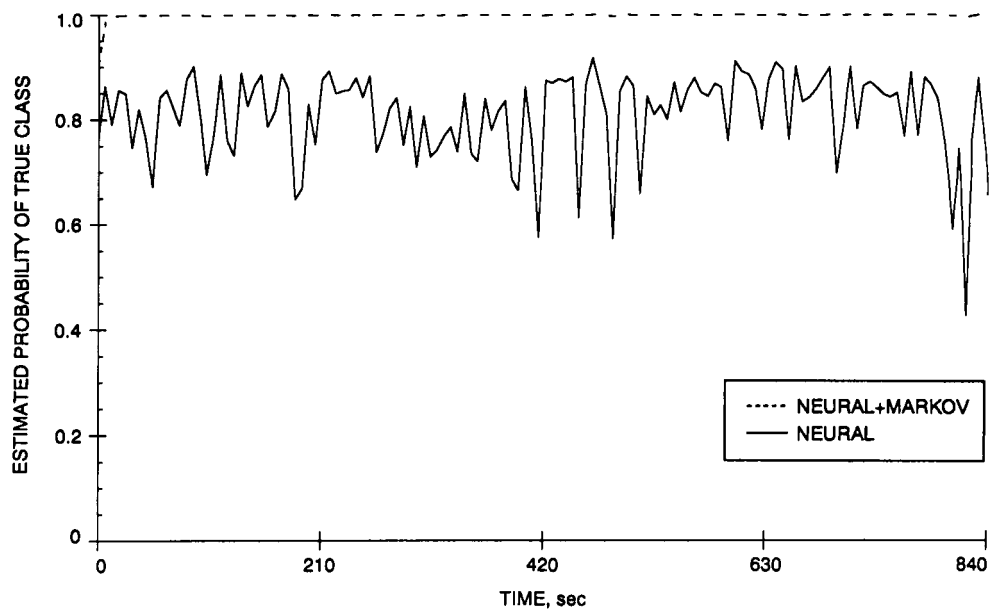


Fig. 2. Comparison of Markov and non-Markov models: estimate of probability of true class (normal conditions) as a function of time for day 316.

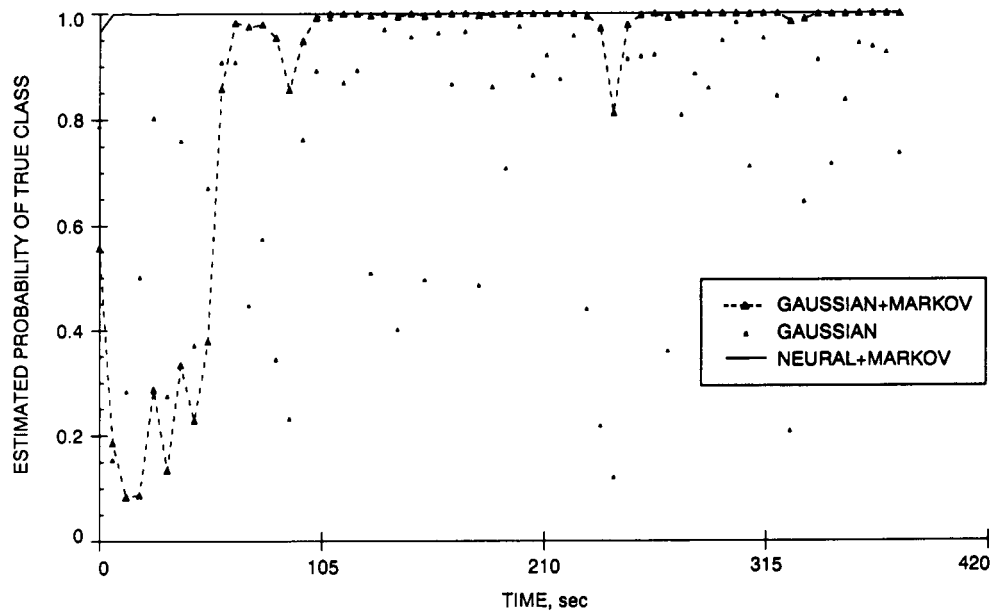


Fig. 3. Comparison of Gaussian-Markov and neural-Markov models: estimate of probability of true class (compensation loss) as a function of time for day 311.

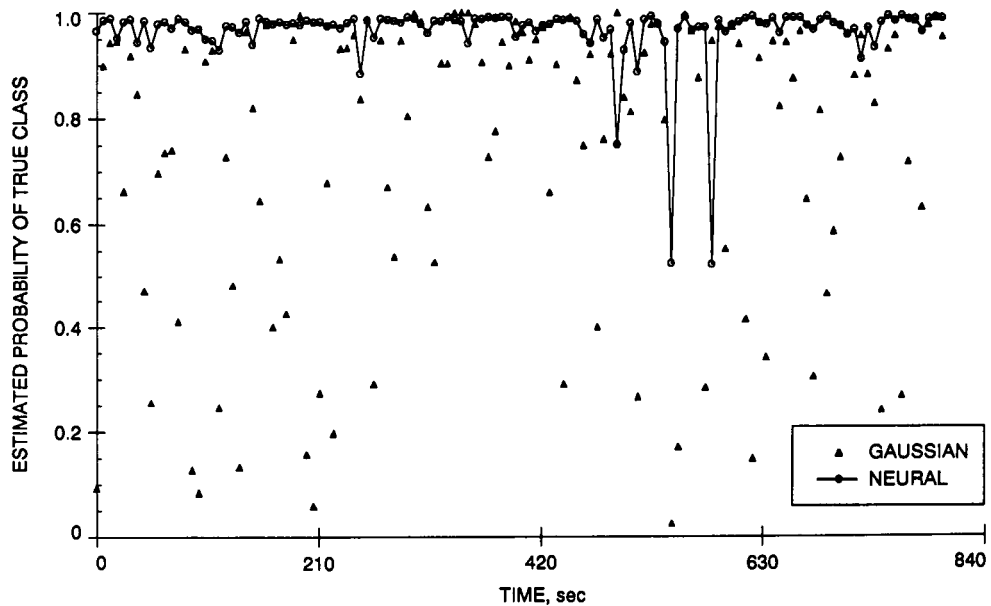


Fig. 4. Comparison of Gaussian and neural models: estimate of probability of true class (compensation loss) as a function of time for day 316.

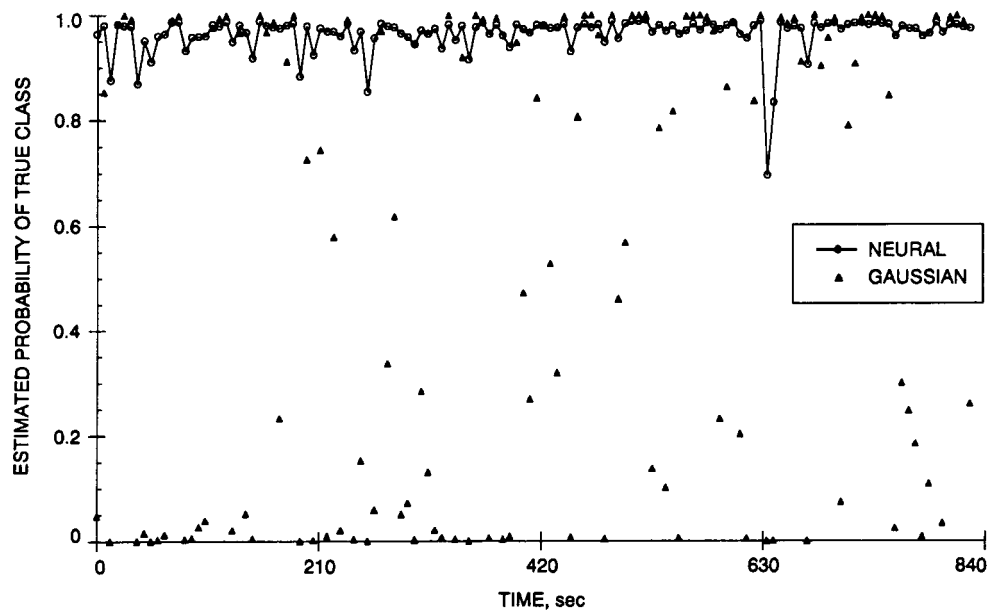


Fig. 5. Comparison of Gaussian and neural models: estimate of probability of true class (tachometer failure) as a function of time for day 316.

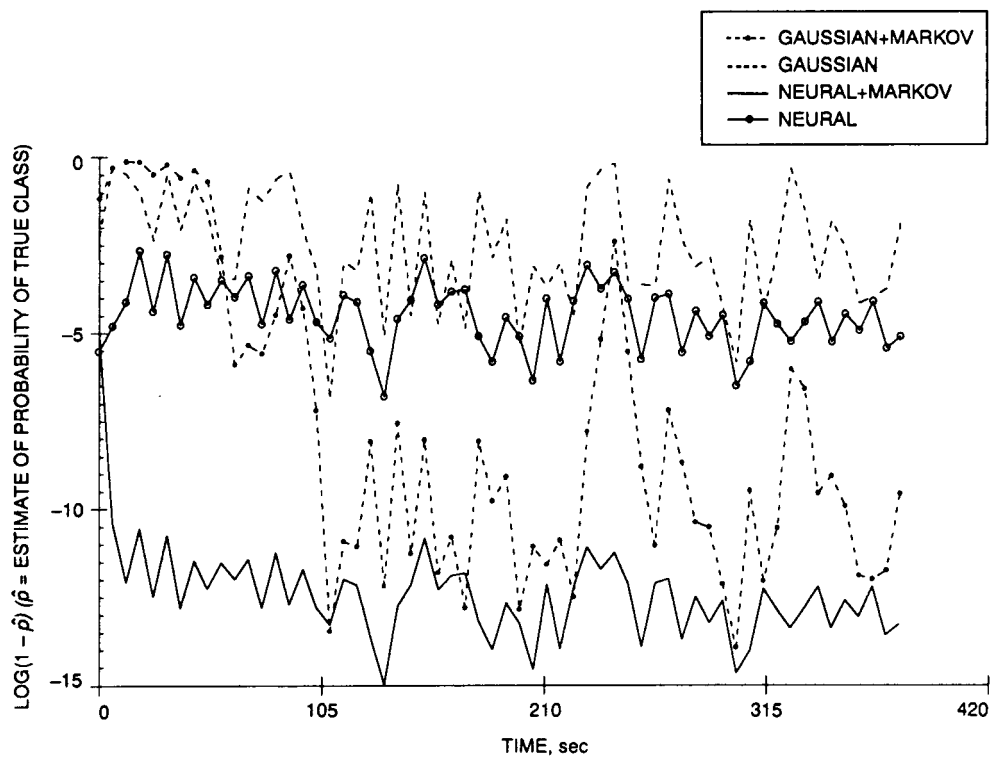


Fig. 6. Comparison of Gaussian-Markov and neural-Markov models: plot of $\log(1 - \hat{p})$, where \hat{p} is an estimate of the probability of the true class (compensation loss) as a function of time for day 311.

Appendix A

Hidden Markov Models

Assume that time t is discretized and the system of interest is always in one of n states. Further assume the system of interest is described by a Markov model, i.e., at each time t the probability that the system is in any state i , $p(s_i(t))$, is only a function of the state j that the system was in at time $t - 1$. Strictly speaking, this is a *first-order* Markov model. Given that the system was in state j at time $t - 1$, the probability $a_{ij} = p(s_i(t)|s_j(t - 1))$ is the *state transition probability* and the $n \times n$ matrix \mathbf{A} with components a_{ij} characterizes the Markov model.

In practice, however, the states may not be directly observable, giving rise to the notion of a *hidden* Markov model. Instead, each state produces an observable set of symptoms or features $\mathbf{x}(t)$ at each time t . The features $\mathbf{x}(t)$ are not a deterministic function of the state, rather they are described by a probability density func-

tion $p(\mathbf{x}(t)|s_i(t))$. Note that the probability of obtaining a particular $\mathbf{x}(t)$ depends on which state i the system is in at time t . It is this fact that makes it possible to infer the probability that the system is in any state i simply by observing the features $\mathbf{x}(t)$.

This model describes quite closely the fault diagnosis problem, namely the system (the antenna elevation-axis servo-control loop) is in some unknown state (normal and failure modes) at each time instant that the classifier looks at the sensor data. The neural network classifier learns the *instantaneous* relationship between states and features, i.e., it provides an estimate of $p(s_i(t)|\mathbf{x}(t))$. However, the best estimate of the state at time t uses all of the information available up to time t , namely the observed feature vectors $\{\mathbf{x}(t), \dots, \mathbf{x}(1)\}$. Hence, the best estimate of state i requires the calculation of $p(s_i(t)|\{\mathbf{x}(t), \dots, \mathbf{x}(1)\})$, which is now derived:

$$\begin{aligned} p(s_i(t)|\{\mathbf{x}(t), \dots, \mathbf{x}(1)\}) &= \frac{p(s_i(t), \{\mathbf{x}(t), \dots, \mathbf{x}(1)\})}{p(\{\mathbf{x}(t), \dots, \mathbf{x}(1)\})} \\ &= \frac{1}{p(\mathbf{X})} \alpha_i(t) \end{aligned}$$

where $\mathbf{X} = \{\mathbf{x}(t), \dots, \mathbf{x}(1)\}$ and $\alpha_i(t) = p(s_i(t), \{\mathbf{x}(t), \dots, \mathbf{x}(1)\})$

$$\begin{aligned} &= \frac{1}{p(\mathbf{X})} \sum_{j=1}^n p(s_i(t), \mathbf{x}(t), \dots, \mathbf{x}(1), s_j(t - 1)) \\ &= \frac{1}{p(\mathbf{X})} \sum_{j=1}^n p(s_i(t), \mathbf{x}(t)|\mathbf{x}(t - 1), \dots, \mathbf{x}(1), s_j(t - 1)) \\ &\quad \times p(\mathbf{x}(t - 1), \dots, \mathbf{x}(1), s_j(t - 1)) \\ &= \frac{1}{p(\mathbf{X})} \sum_{j=1}^n p(s_i(t), \mathbf{x}(t)|s_j(t - 1)) \alpha_j(t - 1) \end{aligned}$$

by the Markov assumption and the definition of α_j

$$= \frac{1}{p(\mathbf{X})} \sum_{j=1}^n p(\mathbf{x}(t)|s_i(t)) p(s_i(t)|s_j(t - 1)) \alpha_j(t - 1)$$

since $\mathbf{x}(t)$ is assumed independent of past states

$$= \frac{1}{p(\mathbf{X})} \sum_{j=1}^n \frac{p(s_i(t)|\mathbf{x}(t))p(\mathbf{x}(t))}{p(s_i(t))} a_{ij} \alpha_j(t-1)$$

by Bayes' rule and the definition of a_{ij}

$$= \frac{p(\mathbf{x}(t))}{p(\mathbf{X})} \sum_{j=1}^n \frac{\hat{p}_i(t)}{p(s_i(t))} a_{ij} \alpha_j(t-1)$$

where $\hat{p}_i(t)$ is the instantaneous probability estimate of state i given $\mathbf{x}(t)$ (namely, the normalized output of the neural network in this case), and $p(s_i(t))$ is the prior probability of state i . Because (by definition)

$$\frac{1}{p(\mathbf{X})} \sum_{i=1}^n \alpha_i(t) = 1$$

$p(\mathbf{x}(t))/p(\mathbf{X})$ can be treated as a constant and factored out. Hence, by virtue of the recursion relation defined

on the α 's above and given $p(s(0))$, the probability of each state at time t can be computed recursively from earlier state information. These basic recursion relations are known as the *forward-backward* relations in the literature. For a more extensive discussion on hidden Markov models and their applications, the reader is referred to [3].

For the experiments described in this article, the initial states $p(s_i(0))$ were set equal to the priors, which in turn were taken to be $1/n = 0.25$. The a_{ij} were set to 0.99 for $i = j$ and 0.01/3 for $i \neq j$.

Appendix B

A Maximum-Likelihood Gaussian Classifier

Consider that there are n classes ω_i , $1 \leq i \leq n$. In turn, the features are described by a d -component feature vector \mathbf{x} . For a Gaussian classifier, one assumes that the probability density $p(\mathbf{x}|\omega_i)$ is multivariate normal for each class, i.e.,

$$p(\mathbf{x}|\omega_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|^{1/2}} e^{(-1/2)(\mathbf{x}-\mu_i)' \Sigma_i^{-1} (\mathbf{x}-\mu_i)}$$

where μ_i is a d -component mean vector and Σ_i is a $d \times d$ covariance matrix.

By Bayes' rule,

$$p(\omega_i|\mathbf{x}) \propto p(\mathbf{x}|\omega_i)p(\omega_i)$$

where $p(\omega_i)$ is the prior probability of class i . Given a particular feature vector \mathbf{x} , one calculates $\gamma_i = p(\mathbf{x}|\omega_i)p(\omega_i)$, from which one gets

$$\hat{p}(\omega_i|\mathbf{x}) = \frac{\gamma_i}{\sum_{j=1}^m \gamma_j}$$

which is the posterior estimate of the probability of class i given the feature data. For the results reported in this article, Σ_i was assumed to be diagonal since there were not sufficient data to obtain reliable estimates of all the elements of the full $d \times d$ ($d = 12$) covariance matrix for each class. The diagonal variance terms σ_i and the means μ_i in each dimension were estimated directly from the data using maximum likelihood estimates. The priors $p(\omega_i)$ were chosen to be $1/n = 0.25$ for each class.

References

- [1] P. Smyth and J. Mellstrom, "Initial Results on Fault Diagnosis of DSN Antenna Control Assemblies Using Pattern Recognition Techniques," *TDA Progress Report 42-101*, vol. January-March 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 136-151, May 15, 1990.
- [2] J. Mellstrom and P. Smyth, "Pattern Recognition Techniques Applied to Performance Monitoring of the DSS 13 34-Meter Antenna Control Assembly," *TDA Progress Report 42-106*, vol. April-June 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 30-51, August 15, 1991.
- [3] A. B. Poritz, "Hidden Markov Models: a Guided Tour," *Proceedings of the International Conference on Acoustics, Speech and Signal Processing 1988*, IEEE Press, New York, pp. 7-13, 1988.

03-32
770
P-26
N92-24311JJ5
CH 508845

Pointing-Error Simulations of the DSS-13 Antenna Due to Wind Disturbances

W. Gawronski

Ground Antennas and Facilities Engineering Section

B. Bienkiewicz

Civil Engineering Department

Colorado State University

R. E. Hill¹

Accurate spacecraft tracking by the NASA DSN antennas must be assured during changing weather conditions. Wind disturbances are the main source of tracking errors. This article presents the development of a wind-force model and simulations of wind-induced pointing errors of DSN antennas. The antenna model includes the antenna structure, the elevation and azimuth servos, and the tracking controller. Simulation results show that pointing errors due to wind gusts are of the same order as errors due to static wind pressure and that these errors (similar to those of static wind pressure) satisfy the velocity quadratic law. The methodology presented is used for wind-disturbance estimation and for design of an antenna controller with wind-disturbance rejection properties.

I. Introduction

Reliable estimates of the pointing error due to wind disturbances are required for operating antennas as well as for designing effective controllers. There are two types of wind-induced pointing errors: steady-state error due to static wind pressure and dynamic error due to wind turbulence. Steady-state error models of antennas have been

investigated by Katow and McGinness [1,2]; the dynamic error of the DSS-14 antenna has been analyzed by Mas-soudi [3]; and the 100-m NRAO antenna has been analyzed by Hill.²

² R. Hill, "Servo Design Studies and Pointing Performance Analysis for a 100-m Aperture Radiotelescope," in *Servo Design Specification for the 100-meter Unblocked Aperture Telescope for the National Radio Astronomy Observatory*, JPL D-7477 (internal document), Jet Propulsion Laboratory, Pasadena, California, September 1990.

¹ Independent consultant to the Ground Antennas and Facilities Engineering Section.

In this article, a recently developed antenna model [4] has been used to analyze wind disturbances. The structural finite-element model, as a part of the antenna model, includes modes of up to 10 Hz. The dynamic wind pressure consists of the spatial pressure on the antenna structure and the wind-force time history. In the new approach presented in this article, the wind force is modeled as a pressure on the antenna dish rather than as an equivalent elevation torque commonly used as a replacement for wind action. Wind-pressure distribution is obtained from the JPL/IDEAS finite-element code,³ based on wind tunnel measurements [5].^{4,5} The wind-force time history is another new addition. It is a linear filter that shapes the wind-force time profile obtained from the Davenport wind-velocity spectrum [6].

The results obtained consist of estimates of pointing error due to static wind pressure and wind gusts for 60- and 90-deg antenna-elevation angles. Results show that the pointing error due to wind turbulence is of the same order as the error due to static pressure, thus it must not be neglected.

II. Wind-Velocity and Wind-Force Models

The wind velocity v is a combination of a steady-state flow (mean velocity, or a quasi-steady component, see [10]) v_m , and a turbulence (gust) v_t

$$v = v_m + v_t \quad (1)$$

The component v_t is a random process with zero mean and standard deviation σ_v . The standard deviation and other wind parameters are determined from the Davenport model [5,7]. In this model, the wind-velocity spectrum S_v for the wind at a height of 10 m is as follows:

$$S_v(f) = 4v_s^2 f^{-1} \frac{x^2}{(1+x^2)^{4/3}} \quad (2)$$

where f is the frequency in hertz; v_s is the shear velocity of the flow in meters per second, see [7]; x is the dimensionless frequency, $x = 1,200f/v_m$; and v_m in meters per second is the mean velocity at 10 m. Denoting $\kappa = (v_s/v_m)^2$ as the surface drag coefficient [Eq. (2)] gives

$$S_v(f) = 4,800v_m \left(\frac{\kappa x}{(1+x^2)^{4/3}} \right) \quad (3)$$

The surface drag coefficient is obtained from the roughness of the terrain, see [7]

$$\kappa = (2.5 \ln(z/z_o))^{-2} \quad (4)$$

where z is the distance from the ground ($z = 10$ m) and z_o is the roughness length (5–20 cm in the DSS-13 environment). For the roughness $z_o = 0.05$ m, one obtains $\kappa = 0.0057$; and for $z_o = 0.2$ m, $\kappa = 0.0105$, which agrees with estimates by Levy and McGinness [7] of κ at the antenna site in Goldstone, California.

It is well known (see [8]) that the standard deviation of a stationary process can be determined from its spectrum, namely

$$\sigma_v = \left(T_v(\infty) - T_v(0) \right)^{1/2} \quad (5a)$$

where

$$T_v(f) = \int_0^f S_v(v) dv = 6\kappa v_m^2 \left(1 - (1+x)^{-1/3} \right) \quad (5b)$$

thus,

$$\sigma_v = \alpha v_m, \quad \alpha = (6\kappa)^{1/2} \quad (6)$$

For a typical surface drag coefficient $\kappa = 0.005$ – 0.010 , one obtains $\alpha = 0.17$ – 0.25 .

Although the force acts on the whole antenna structure, in this article the wind force is assumed to act only on the antenna dish. This is a fair assumption, since most of the wind energy is absorbed by the tipping structure. The force distribution on the dish surface yields important information for determining the wind action on the antenna. By using the available quasi-static wind-tunnel

³ R. Levy and D. Strain, JPL-IDEAS Finite Element Analysis and Design Optimization Program, JPL NPO-17783 (internal document), Jet Propulsion Laboratory, Pasadena, California, October 1988.

⁴ R. B. Blaylock, "Aerodynamic Coefficients for a Model of a Paraboloidal Reflector Directional Antenna Proposed for a JPL Advanced Antenna System," JPL Memorandum CP-6 (internal document), Jet Propulsion Laboratory, Pasadena, California, May 1, 1964.

⁵ N. L. Fox, "Load Distributions on the Surface of Paraboloidal Reflector Antennas," JPL Memorandum CP-4 (internal document), Jet Propulsion Laboratory, Pasadena, California, July 1962.

data,^{6,7} the spatial force distribution can be obtained from the JPL-IDEAS model.⁸ As a result, the wind force as a time and spatial variable becomes a time variable in the dynamic model. In the antenna model, the force acting on the dish has two components: F_x and F_y (wind from x and y directions). Since the time characteristics of the wind force do not depend on its direction, it is not necessary to distinguish the direction, thus the wind force is simply denoted by F .

Statistical properties of the wind force are determined from the previously described wind-velocity parameters. At a given height, the force is proportional to the square of the velocity [9]

$$F = kv^2 \quad (7)$$

and the velocity consists of a constant mean flow with a superimposed fluctuation, as in Eq. (1). Therefore, the force is decomposed into a steady force F_m and a turbulent force F_t with zero mean value

$$F = F_m + F_t \quad (8)$$

The static force F_m (lb) is determined for the speed v_m (mph)

$$F_m = F(v_m) = kv_m^2 = \alpha_w^2 F_{om} \quad (9)$$

where $\alpha_w = v_m/100$, $F_{om} = 1$ lb is the force at a wind velocity of 100 mph, and $k = 0.0001F_{om}$. The gust F_t , on the other hand, is determined as follows:

$$F_t = \alpha_f v_t \quad (10)$$

where $\alpha_f = \partial F / \partial v|_{v=v_m} = 2kv_m = 2F_m/v_m$ depends on the wind velocity v_m . From Eq. (10) the wind-force spectrum can be obtained as follows:

$$S_f(f) = \alpha_f^2 S_v(f) \quad (11)$$

Thus, the standard deviation σ_f of the gust is

$$\sigma_f = |\alpha_f| \sigma_v \quad (12)$$

⁶ R. B. Blaylock, op. cit.

⁷ N. L. Fox, op. cit.

⁸ R. Levy and D. Strain, op. cit.

Finally, by introducing Eq. (6) to Eq. (12), one obtains

$$\sigma_f = 2\alpha F_m \quad (13)$$

for the static force F_m given by Eq. (9).

III. Wind-Force Simulations

Static and dynamic wind forces have been simulated. The static force is determined from Eq. (9), and the static force simulations are obvious. The dynamic force is generated by a filter with a transfer function $T(f)$

$$T(f) = \frac{a_0 + a_1 s + a_2 s^2 + \dots + a_m s^m}{b_0 + b_1 s + b_2 s^2 + \dots + b_m s^m} \quad (14)$$

The parameters a_i and b_i , $i = 1, \dots, m$, $m > 0$ are determined so that the power spectrum of $T(f)$ is the best approximation of the wind-power spectrum $S_f(f)$. For this purpose, the square root s_f of the power spectrum S_f is used

$$s_f(f) = (S_f(f))^{1/2} \quad (15)$$

The order m of the filter and the parameters a_i and b_i , $i = 1, \dots, m$ can be determined. For the frequency interval $[f_1, f_2]$, $f_2 > f_1$, denote the norm $\| \cdot \|_w$ of a function $g(f)$ as

$$\|g\|_w = \left(\int_{f_1}^{f_2} w(f) g^2(f) df \right)^{1/2} \quad (16)$$

where $w(f) > 0$, $f \in [f_1, f_2]$ is a weighting function. Denote the poles of the transfer function λ_i , $i = 1, \dots, m$, then the filter parameters, and the filter order should be found so that the index J , defined as follows,

$$J = \|s_f(f) - |T(f)|\|_w \quad (17)$$

is minimal, subject to constraints

$$\text{Re}(\lambda_i) < 0, \quad \text{for } i = 1, \dots, m \quad (18)$$

The constraints assure filter stability.

The plot of the function $s_f(f)$, obtained from Eq. (15) (the Davenport model for 30-mph wind speed), is shown by the solid line in Fig. 1. A linear filter that generates the wind force of a spectrum close to the Davenport model is determined by minimizing J subject to constraints in Eq. (18). The values of the Davenport spectrum for high frequencies (in the case considered for frequencies 1–1,000 Hz) are small; therefore, they should be weighted significantly to obtain a good fit in this frequency range. The following weights have been assigned: $w = 3$ for frequencies 0.1–1 Hz, $w = 10$ for 1–10 Hz, $w = 30$ for 10–100 Hz, and $w = 100$ for 100–1,000 Hz. For different integers m , the index J has been minimized; the integer $m = 4$ was found to be the lowest filter order for which the results have been satisfactory. The filter parameters for the minimal solution are:

$$\begin{aligned} a_0 &= 0.00037, \quad a_1 = -0.92431, \quad a_2 = -138.7468, \\ a_3 &= -56.4776, \quad a_4 = -0.03952 \\ b_0 &= -0.000306, \quad b_1 = -0.18103, \quad b_2 = -8.7769, \\ b_3 &= -57.9502, \quad b_4 = -11.8569 \end{aligned}$$

The magnitude of the filter-transfer function is shown by the dashed line in Fig. 1. The figure shows a good approximation of the Davenport spectrum within the interval of interest [$f_1 = 0.001$, $f_2 = 10$] Hz. The wind force generated by the filter for 30-mph wind is shown in Fig. 2.

A similar approach was used to determine wind filters and wind forces for wind velocities $v_m = 10, 20, 40$, and 50 mph.

IV. Pointing-Error Simulations

The antenna state-space model with the position loop closed, described in [4], is as follows:

$$\dot{x} = Ax + B_c u + B_w F, \quad y = Cx + Du \quad (19)$$

where x is a state vector $n \times 1$; u is the command signal $p \times 1$; F is the wind force (single input); y is the output $q \times 1$. The input matrix B_w is determined so that for a wind of 100 mph, a static wind force $F_{om} = 1$ lb is applied. The plot of drag force versus antenna-elevation angle obtained from the wind-tunnel experiment⁹ is shown in Fig. 3. In this article, two elevation positions of the

antenna are simulated: 90 deg and 120 deg (that is, 60 deg with wind blowing from behind the dish). Note from Fig. 3 that the drag is almost the same for the both positions. Nevertheless, it will be shown later that the effects of the wind actions on the antenna are distinct.

Since the antenna model is linear, the static force (mean wind force) and dynamic force (gust force) are simulated separately. However, they can be superimposed, if necessary. For the static wind disturbances the system with the position loop closed [represented by Eq. (19)] is simulated with the zero command signal u and constant wind disturbance F_m . The steady-state response (elevation and cross-elevation pointing error) is obtained for a 60- and 90-deg antenna-elevation angle. The plots of pointing error due to static wind from the x and y directions for different wind speeds are shown in Fig. 4. The following results for static wind load are obtained: For 30-mph x -direction wind, the elevation pointing error is 9.0 mdeg for a 60-deg elevation angle and 1.7 mdeg for a 90-deg elevation angle. The cross-elevation pointing error is 5.8 mdeg for a 60-deg elevation angle and 6.6 mdeg for a 90-deg elevation angle. For 30-mph y -direction wind, the elevation pointing error is 1.1 mdeg for a 60-deg elevation angle and 1.3 mdeg for a 90-deg elevation angle; the cross-elevation pointing error is 0.13 mdeg for a 60-deg elevation angle and 0.22 mdeg for a 90-deg elevation angle. The small pointing error due to wind in the y direction is a result of the integrating action of the controller.

The gust wind load was applied to the antenna model. The responses of the antenna (elevation and cross-elevation pointing errors, elevation and azimuth encoder readings) are presented in Fig. 5 for x -direction wind and in Fig. 6 for y -direction wind. Wind force was simulated for velocities 10, 20, 30, 40, and 50 mph for 60- and 90-deg elevation angles of the antenna. The results are presented in Fig. 7 for x -direction wind and 60-deg elevation angle (solid line) and 90-deg elevation angle (dashed line). Figure 8 shows the results for y -direction wind. Note that like the results of the static load, the results of the dynamic loading show the pointing errors to be proportional to the square of the velocity (the symbols "o" in Figs. 7 and 8 denote the simulation results, and the lines denote the least-squares quadratic approximation). It can be seen in Fig. 8 that the integrating action of the controller annihilated a significant part of the pointing error due to static y -direction wind; however, the dynamic part of the y -direction wind is not compensated for by the controller, which causes a significant pointing error.

The spectra of the wind-gust response were analyzed. The spectra of the elevation and cross-elevation pointing

⁹ Engineering Data 34-M Antenna, JPL Drawing 9493280 (internal document), Jet Propulsion Laboratory, Pasadena, California, 1988.

errors, and elevation and azimuth encoder readings for x -direction wind gusts are shown in Fig. 9, and in Fig. 10 for y -direction wind gusts. The spectra show that modes of low frequency (2 Hz), medium (4 Hz), and high frequency (9 Hz) are important in wind-gust simulations.

Finally, a power spectral-density analysis of the pointing errors was performed for the antenna at the 90-deg elevation position, and results were compared with the simulation results of Figs. 5, 6, 9, and 10. The spectral analysis provides a precise measure of root-mean-square (rms) error, limited only by step-size effects. It serves as a convenient method for statistical verification of simulations.

The power spectral-density function of the pointing error $S_e(f)$ was calculated according to

$$S_e(f) = W^2(f) S_f(f) \quad (20)$$

where $W(f)$ is the amplitude-versus-frequency response function between the respective wind-disturbance inputs and pointing-error outputs of the system described by Eq. (19), and $S_f(f)$ represents the wind-force spectrum from Eq. (11). The corresponding spectral distributions and rms errors were determined from a numerical integration of $S_e(f)$ over frequency. An integration routine that adopted step size in the vicinity of peaks of $S_e(f)$ was employed to minimize computational errors.

The frequency-response functions for elevation- and cross-elevation pointing errors are plotted in Figs. 11(a) and (b) for x -axis, and in Figs. 12(a) and (b) for y -axis wind direction. The corresponding power spectral density and spectral distribution functions for pointing errors are shown in Fig. 13 for x -axis winds and in Fig. 14 for y -axis winds. The mean wind velocity in both cases was 30 mph. The rms pointing errors were calculated from the corresponding spectral distribution functions at a frequency of 10.0 Hz. For x -axis direction wind the elevation pointing error is 1.60 mdeg and the cross-elevation pointing error is 5.15 mdeg, while for y -axis direction wind, the elevation pointing error is 5.53 mdeg, and the cross-elevation pointing error is 0.65 mdeg. These errors agree well with the values obtained from simulations of Figs. 5 and 6.

V. Conclusions

Wind forces acting on the antenna structure have been modeled in this article. Their static and dynamic parts have been applied as a disturbance to the closed-loop antenna model at 60- and 90-deg elevation angles. The results show that both the static and dynamic parts of the pointing error are of equal importance, and that despite the dominant presence of low-frequency components in the wind-disturbance spectrum, the high-frequency modes of the antenna have been excited, which caused significant pointing errors. In order to improve the analytical results, experimental verification should be the first priority.

Acknowledgments

The authors thank Roy Levy, Douglas Strain, and Ben Parvin for their extensive and useful discussions.

References

- [1] M. S. Katow and H. D. McGinness, *Wind Load Predictions for the 64-Meter-Diameter Antenna*, Technical Report 32-1526, Jet Propulsion Laboratory, Pasadena, California, vol. 15, June 15, 1973.
- [2] M. S. Katow, "Aerodynamic Static Differential Pressure Values for the 50 Percent Porous Reflector Dish," *DSN Progress Report 42-29*, vol. July and August 1975, Jet Propulsion Laboratory, Pasadena, California, pp. 60-65, October 15, 1975.
- [3] M. Massoudi, "Tracking Error of 100-m Antenna due to Wind Gust," *TDA Progress Report 42-48*, vol. September and October, Jet Propulsion Laboratory, Pasadena, California, pp. 94-101, December 15, 1978.
- [4] W. Gawronski and J. A. Mellstrom, "Modeling and Simulations of the DSS 13 Antenna Control System," *TDA Progress Report 42-106*, vol. April-June, Jet Propulsion Laboratory, Pasadena, California, pp. 205-248, August 1991.
- [5] R. Levy and D. Kurtz, *Compilation of Wind Tunnel Coefficients for Parabolic Reflectors*, JPL Publication 78-16, Jet Propulsion Laboratory, Pasadena, California, April 15, 1978.
- [6] A. G. Davenport, "The Spectrum of Horizontal Gustiness Near the Ground in High Winds," *Journal of Royal Meteorol. Society*, vol. 87, pp. 194-211, 1961.
- [7] R. Levy and H. McGinness, *Wind Power Predictions Model*, TM 33-802, Jet Propulsion Laboratory, Pasadena, California, November 1976.
- [8] E. Simiu and R. H. Scanlan, *Wind Effects on Structures*, New York: Wiley Interscience, 1978.
- [9] J. S. Bendat and A. G. Piersol, *Random Data*, New York: Wiley Interscience, 1986.
- [10] J. Vellozzi and E. Cohen, "Gust Response Factors," *Proceedings of the ASCE, Journal of the Structural Division*, vol. 94, no. ST6, pp. 1295-1313, June 1968.

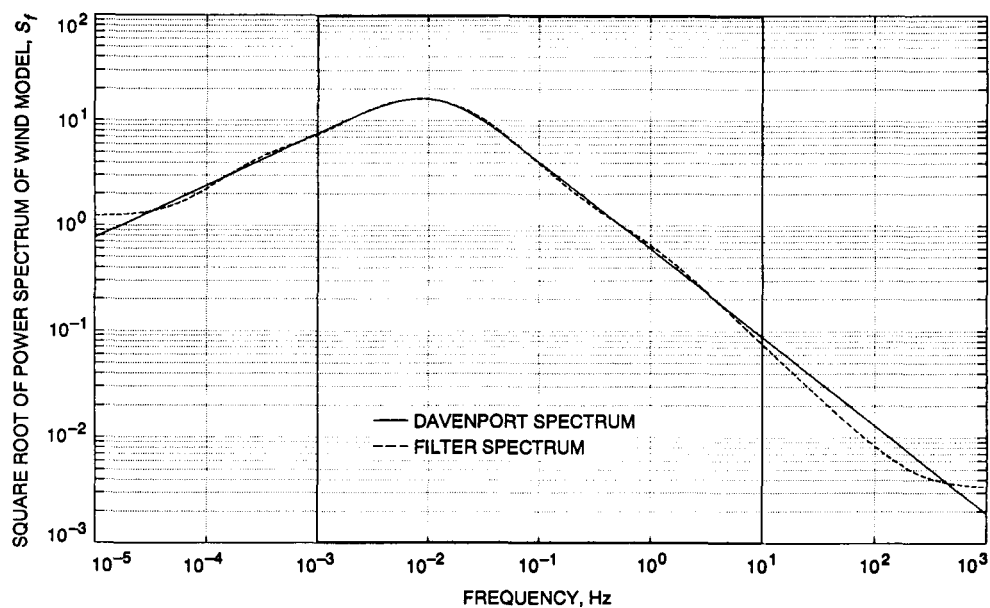


Fig. 1. Square root of power spectrum of the Davenport wind-gust model, for 30-mph wind at the DSS-13 antenna site.

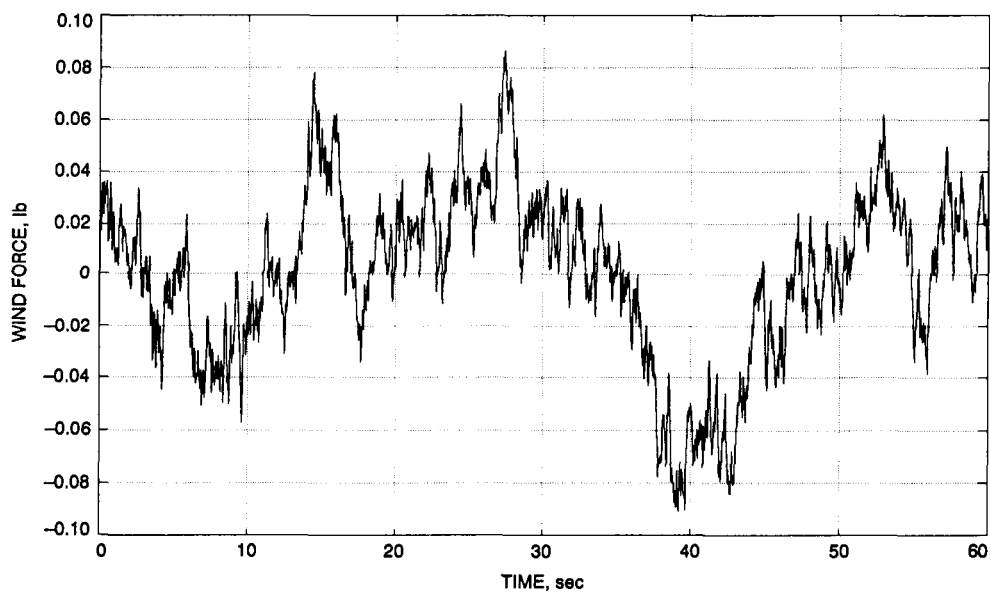


Fig. 2. Wind force generated by the filter for a 30-mph wind gust.

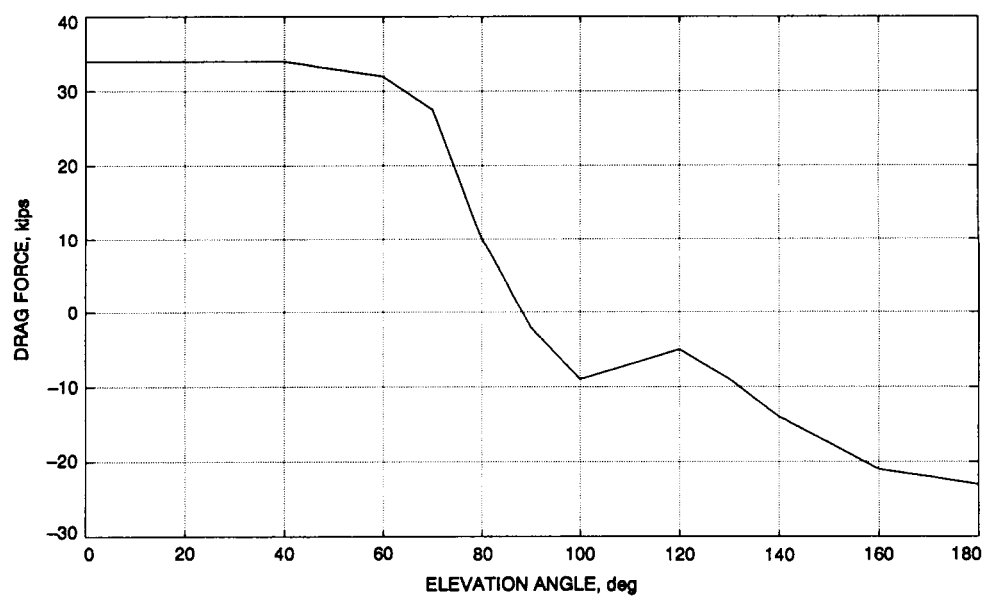


Fig. 3. Experimentally determined wind drag force versus elevation angle.

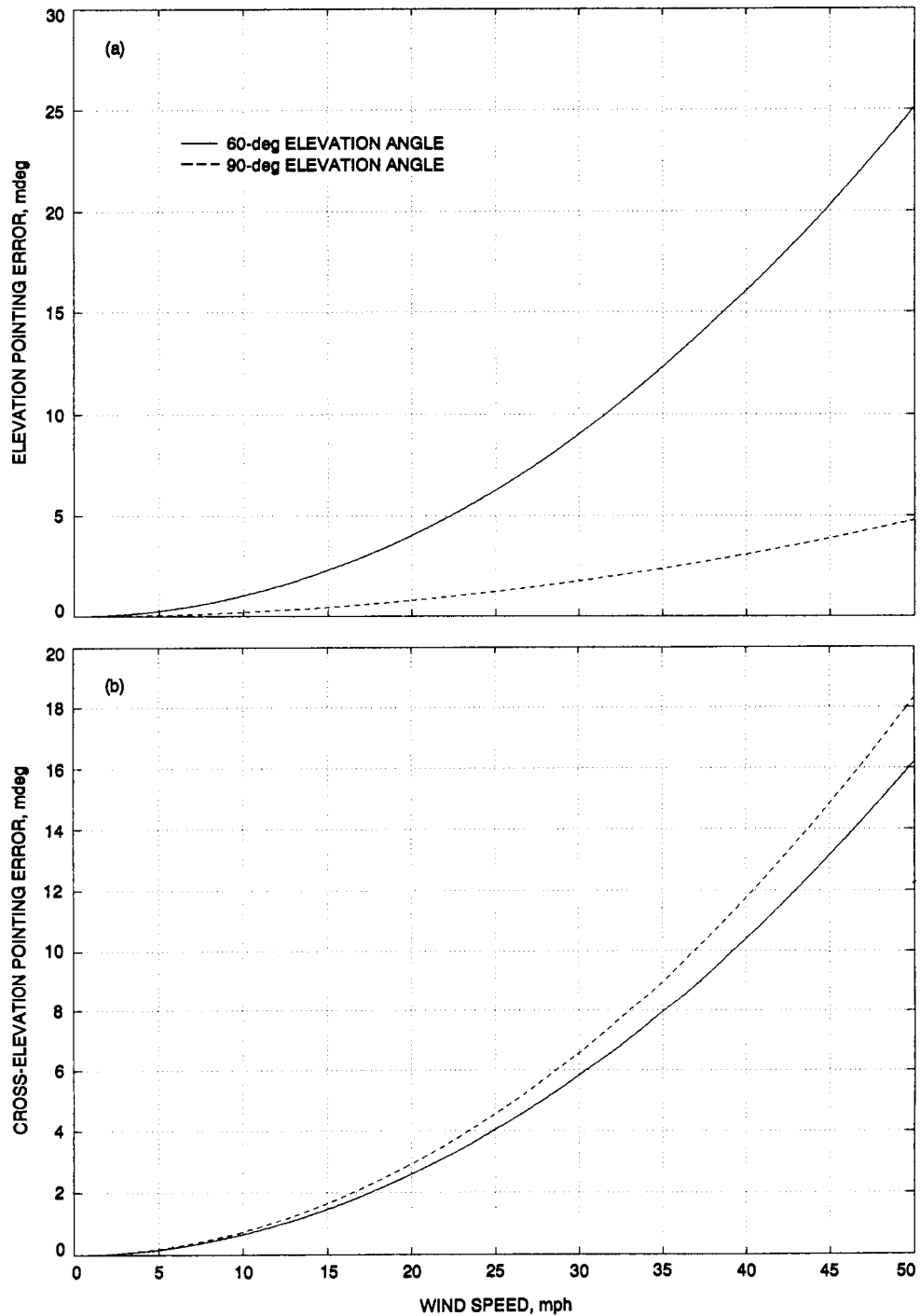


Fig. 4. Pointing error due to static wind from x- and y-direction for different wind speeds: (a) elevation pointing error; (b) cross-elevation pointing error; (c) elevation encoder reading; and (d) azimuth encoder reading.

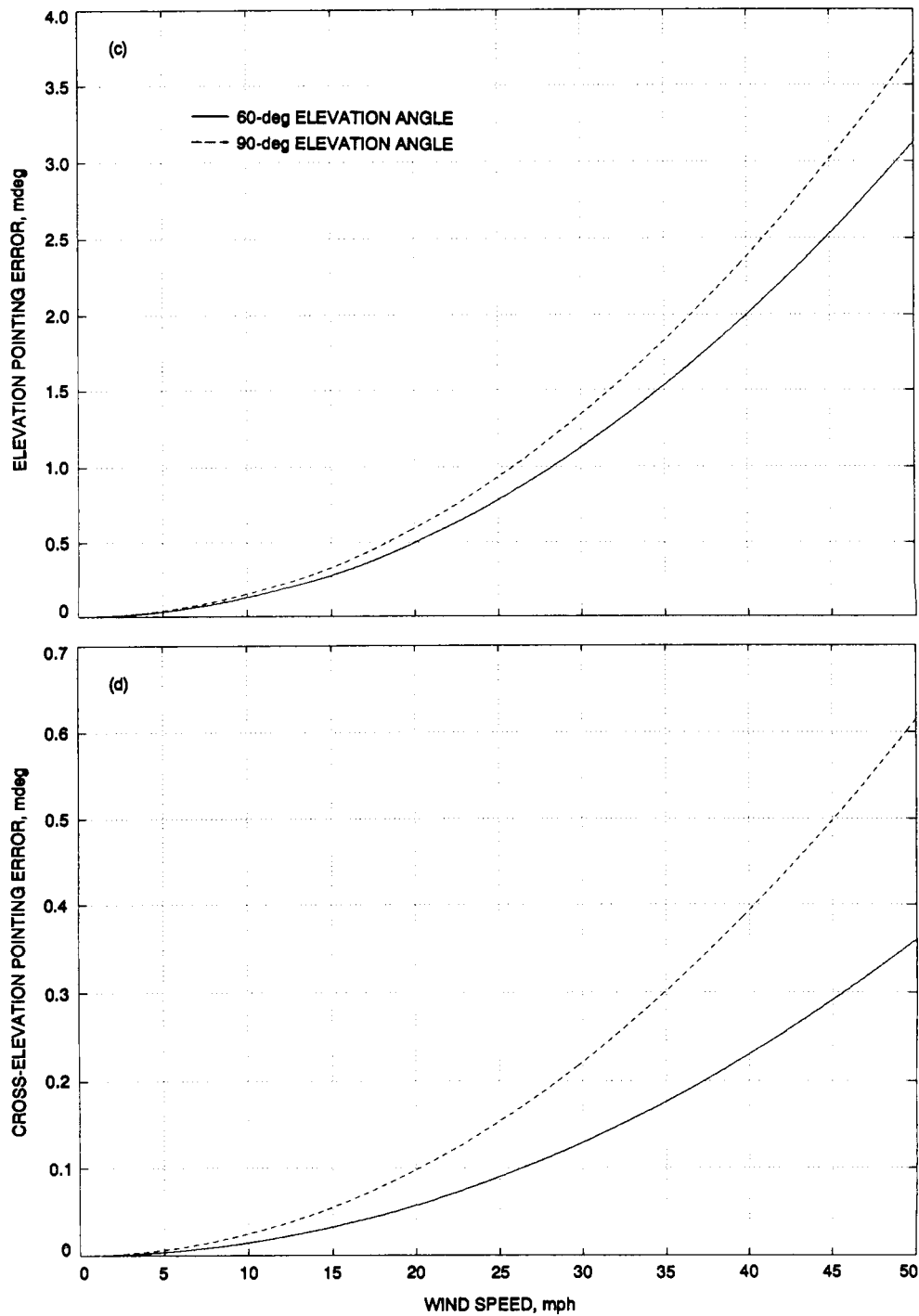


Fig. 4 (contd).

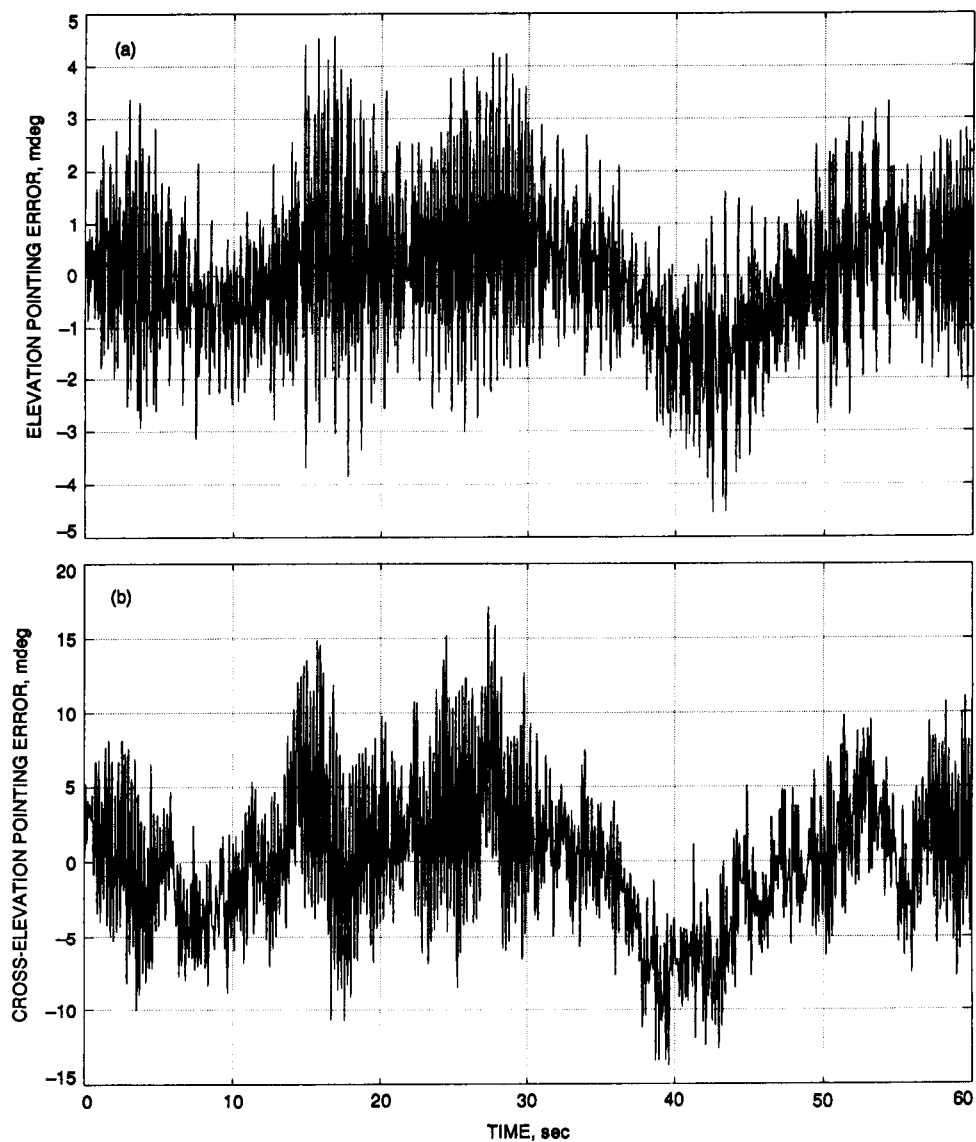


Fig. 5. Antenna responses to x-direction wind: (a) elevation pointing error; (b) cross-elevation pointing error; (c) elevation encoder reading; and (d) azimuth encoder reading.

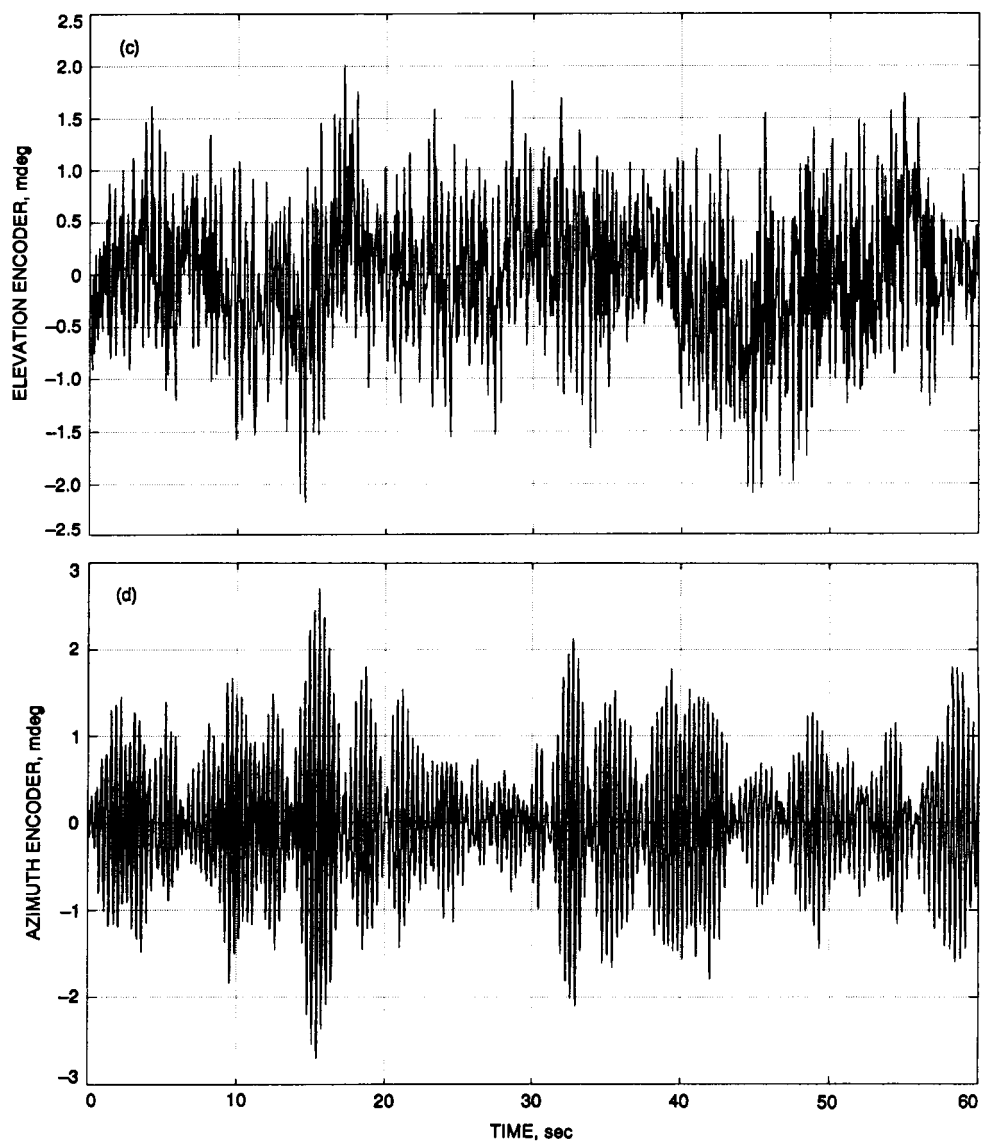


Fig. 5 (contd).

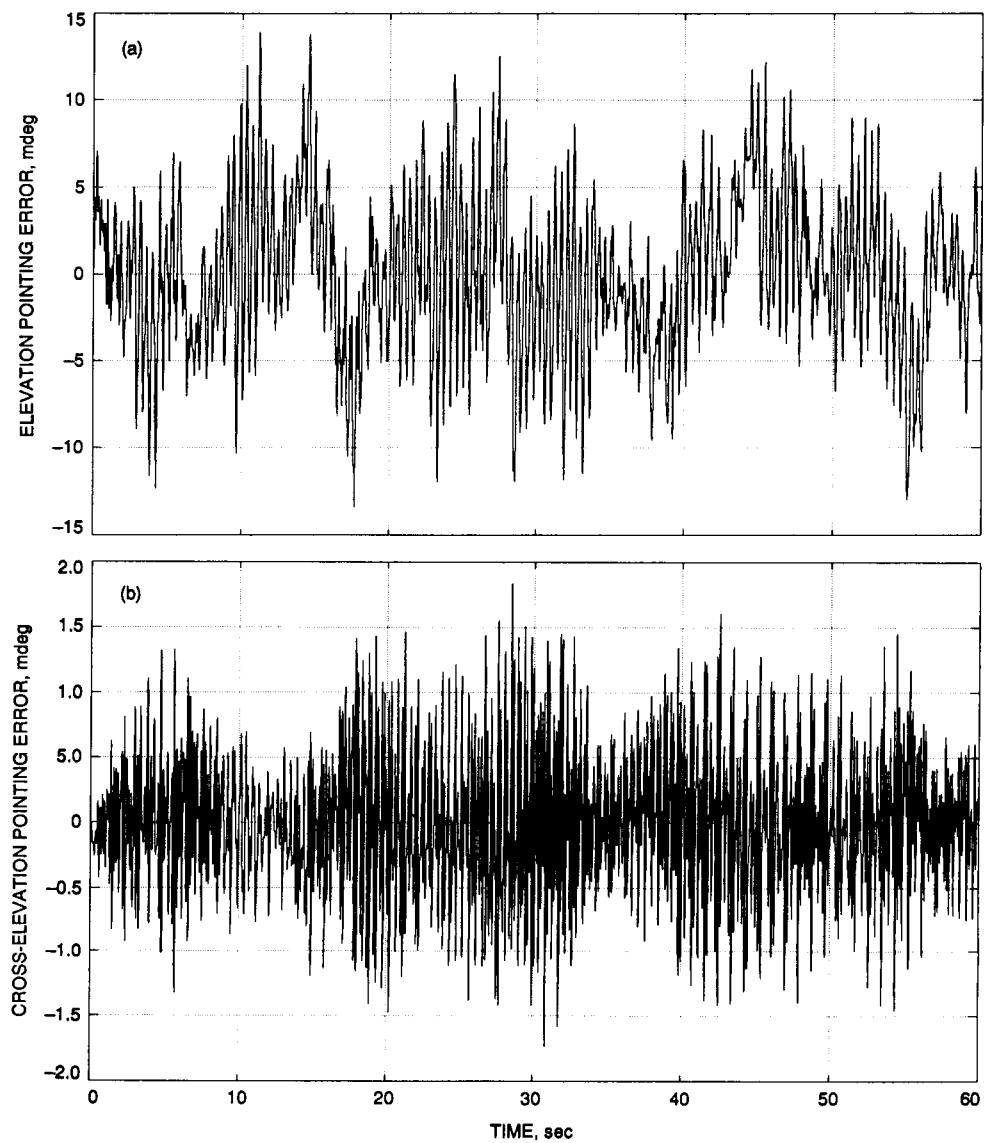


Fig. 6. Antenna responses to y-direction wind: (a) elevation pointing error; (b) cross-elevation pointing error; (c) elevation encoder reading; and (d) azimuth encoder reading.

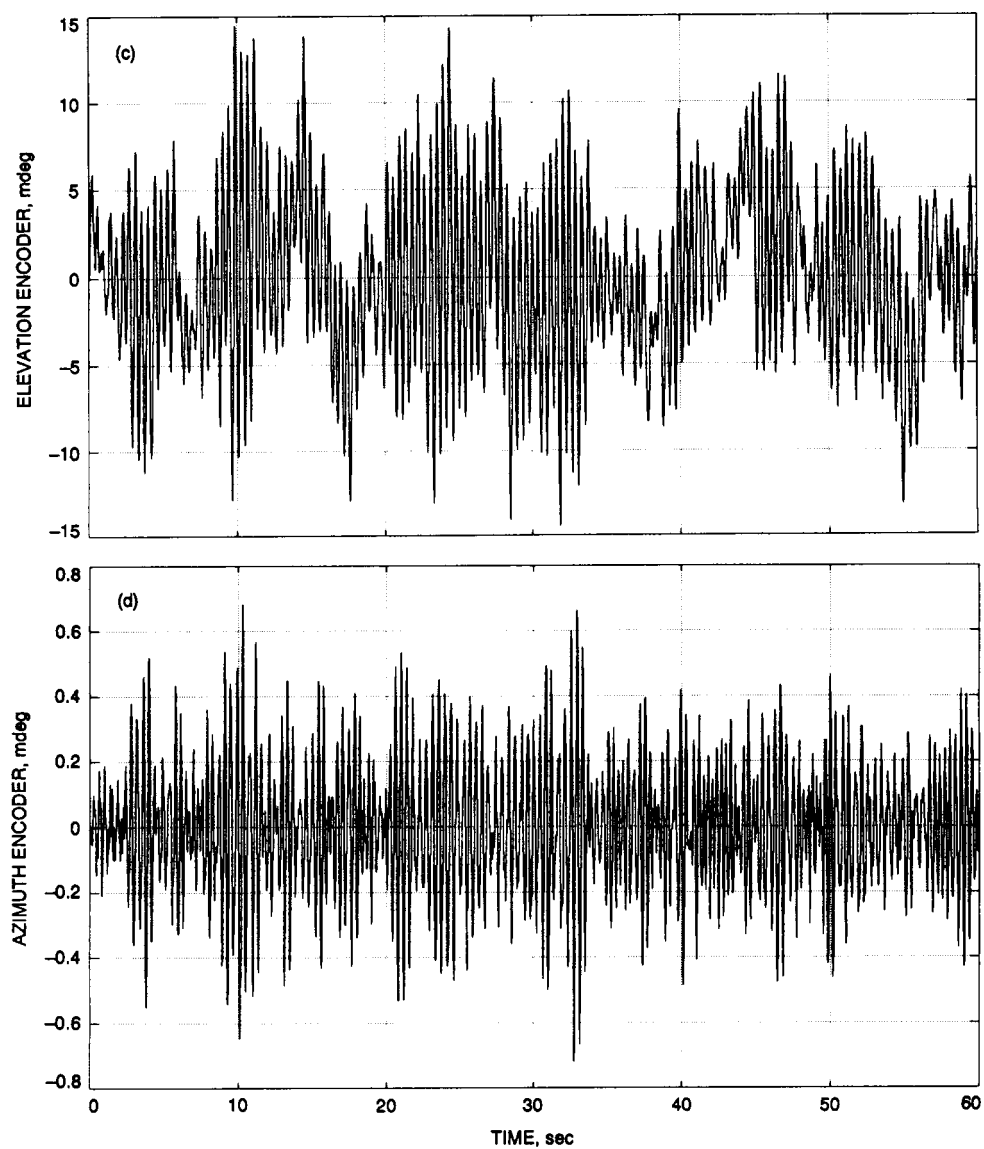


Fig. 6 (contd).

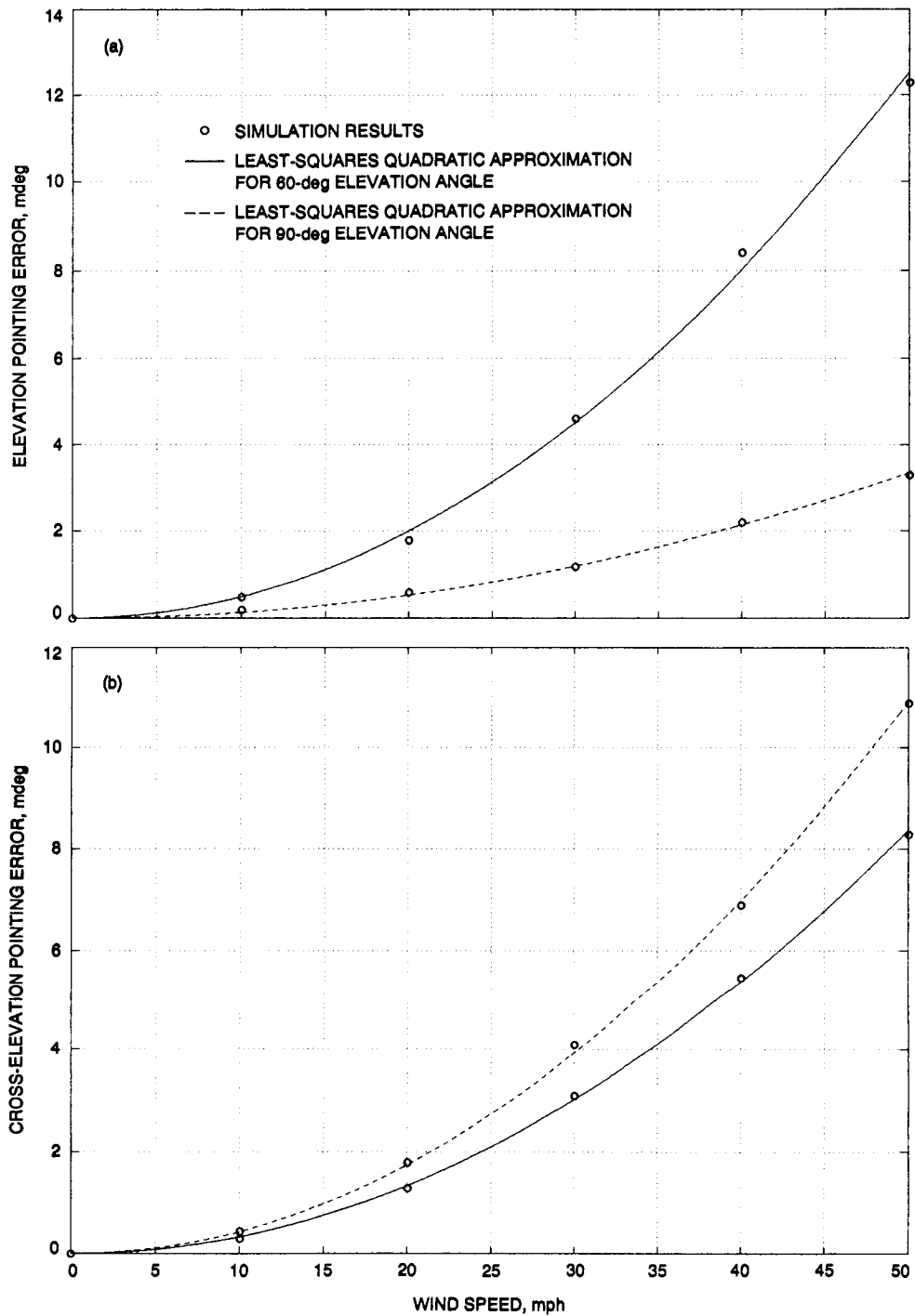


Fig. 7. The quadratic law for x-direction wind gusts: (a) elevation pointing error; (b) cross-elevation pointing error; (c) elevation encoder reading; and (d) azimuth encoder reading.

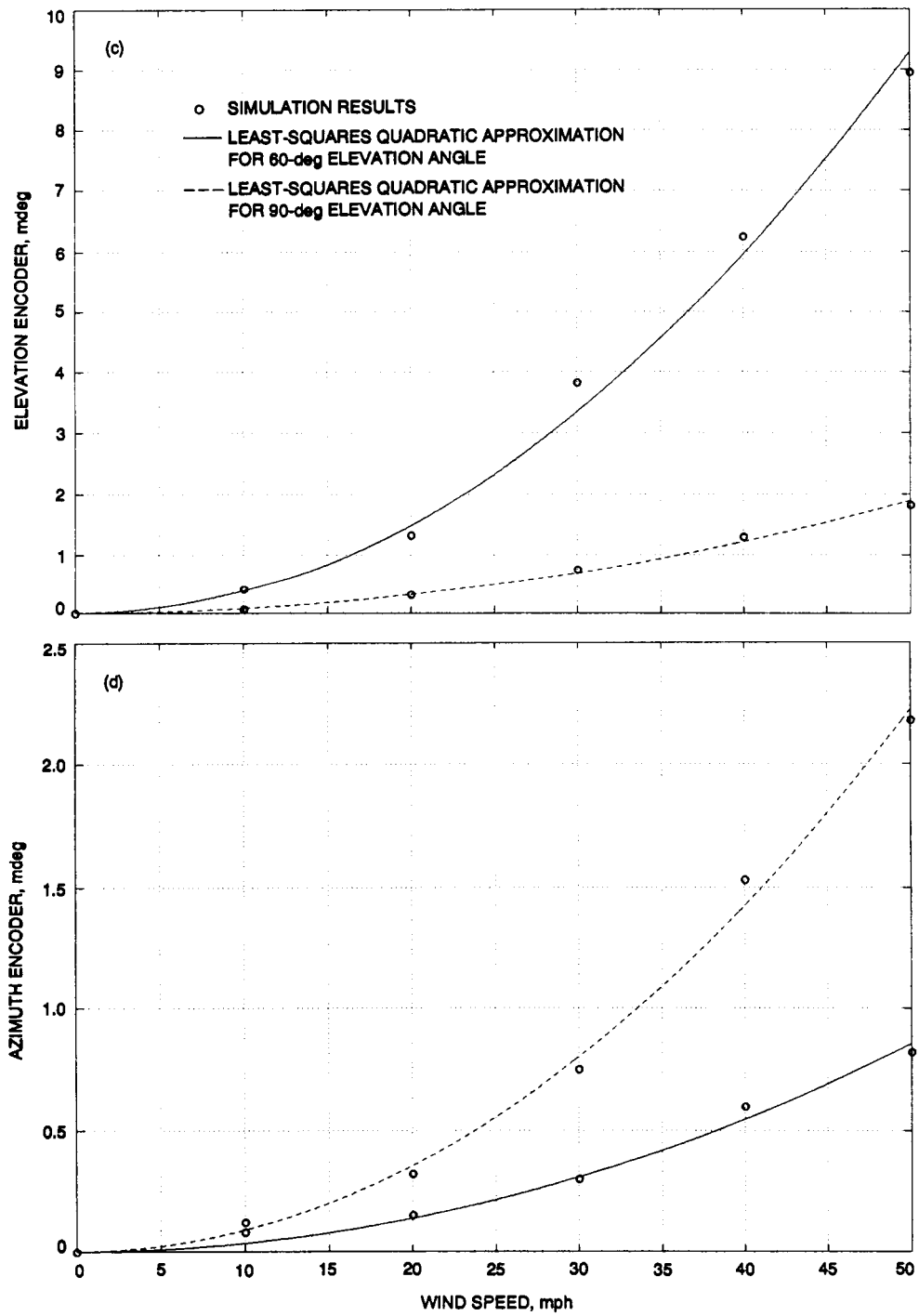


Fig. 7 (contd).

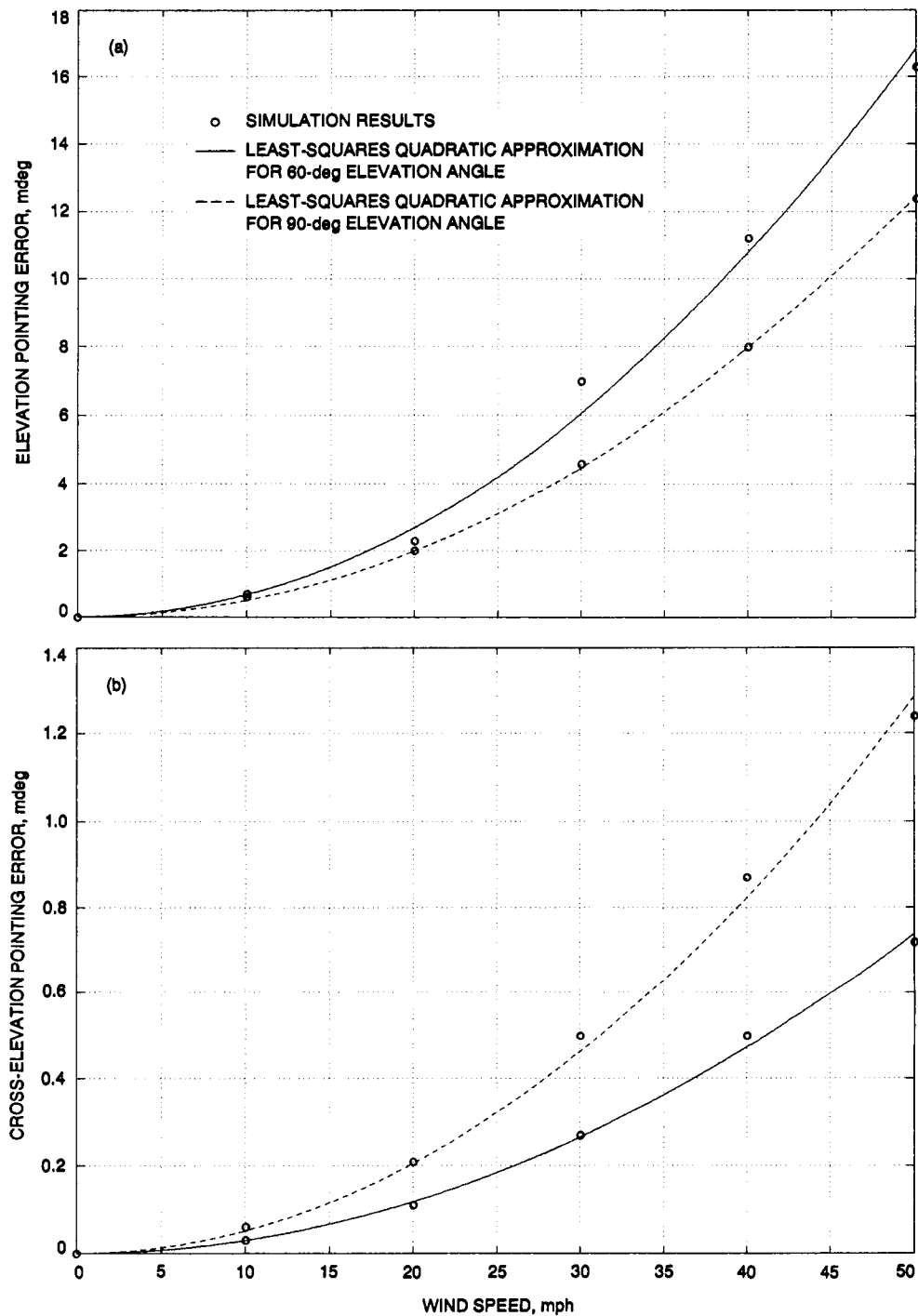


Fig. 8. The quadratic law for y-direction wind gusts: (a) elevation pointing error; (b) cross-elevation pointing error; (c) elevation encoder reading; and (d) azimuth encoder reading.

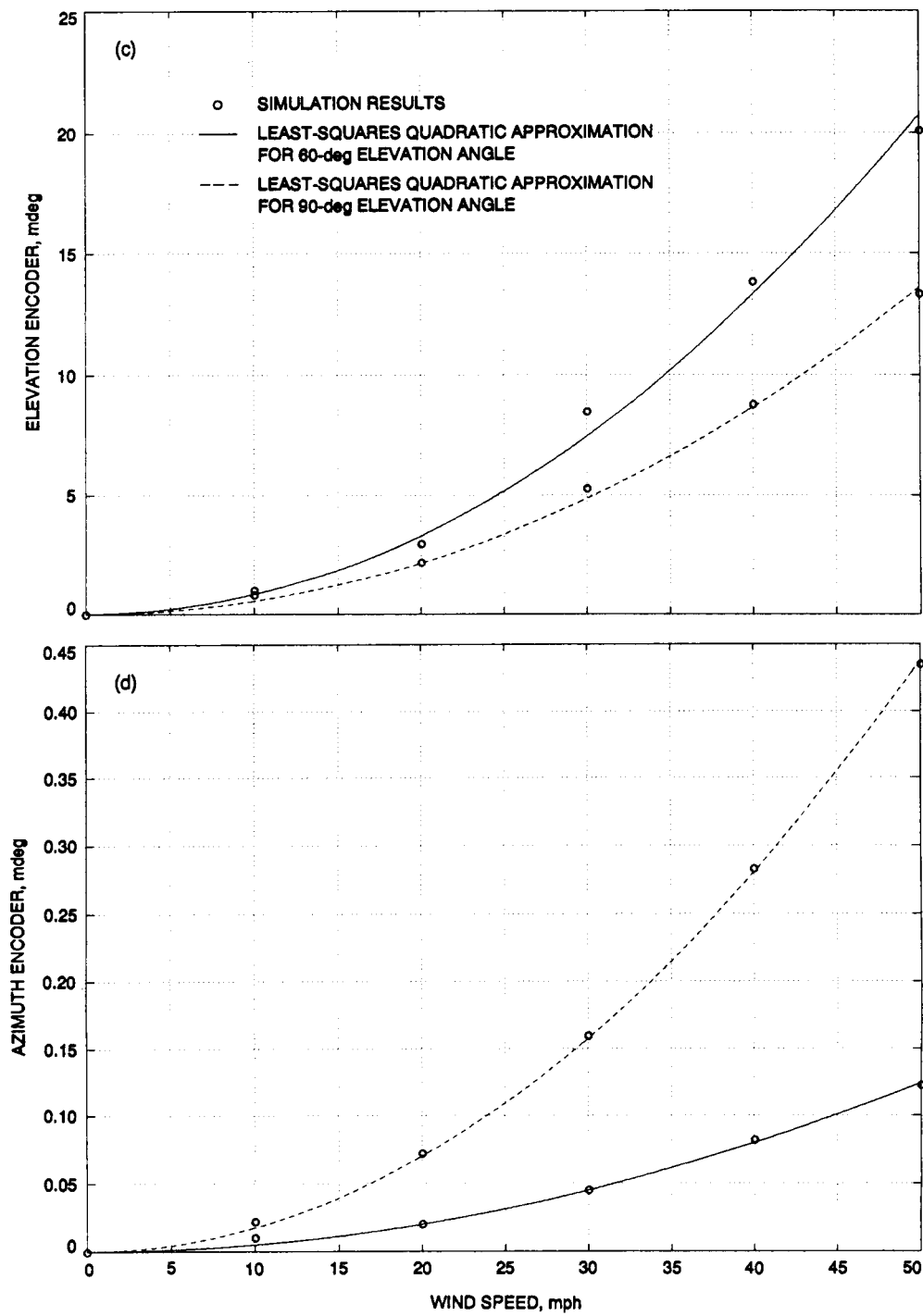


Fig. 8 (contd.).

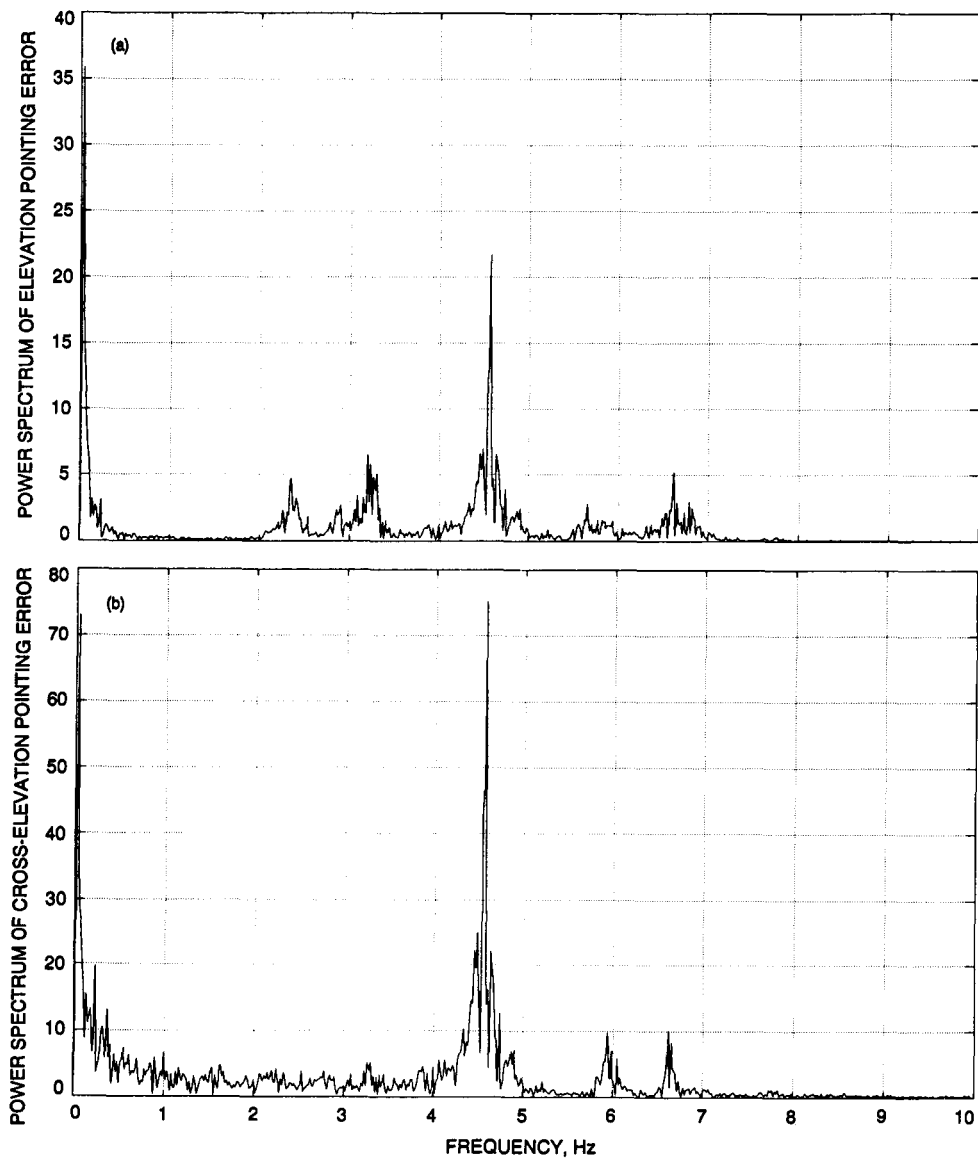


Fig. 9. Spectra for x-direction wind gusts: (a) elevation pointing error; (b) cross-elevation pointing error; (c) elevation encoder reading; and (d) azimuth encoder reading.

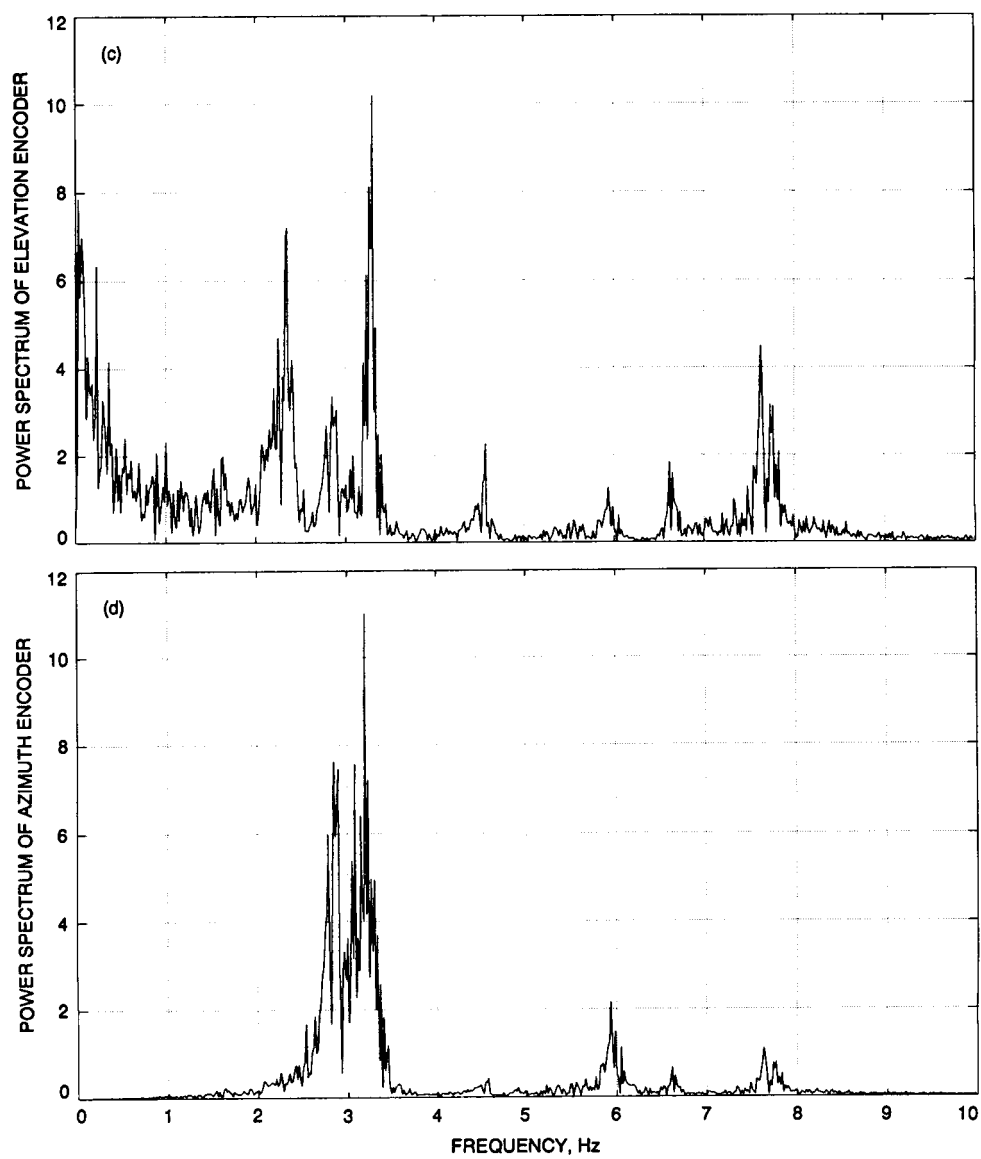


Fig. 9 (contd).

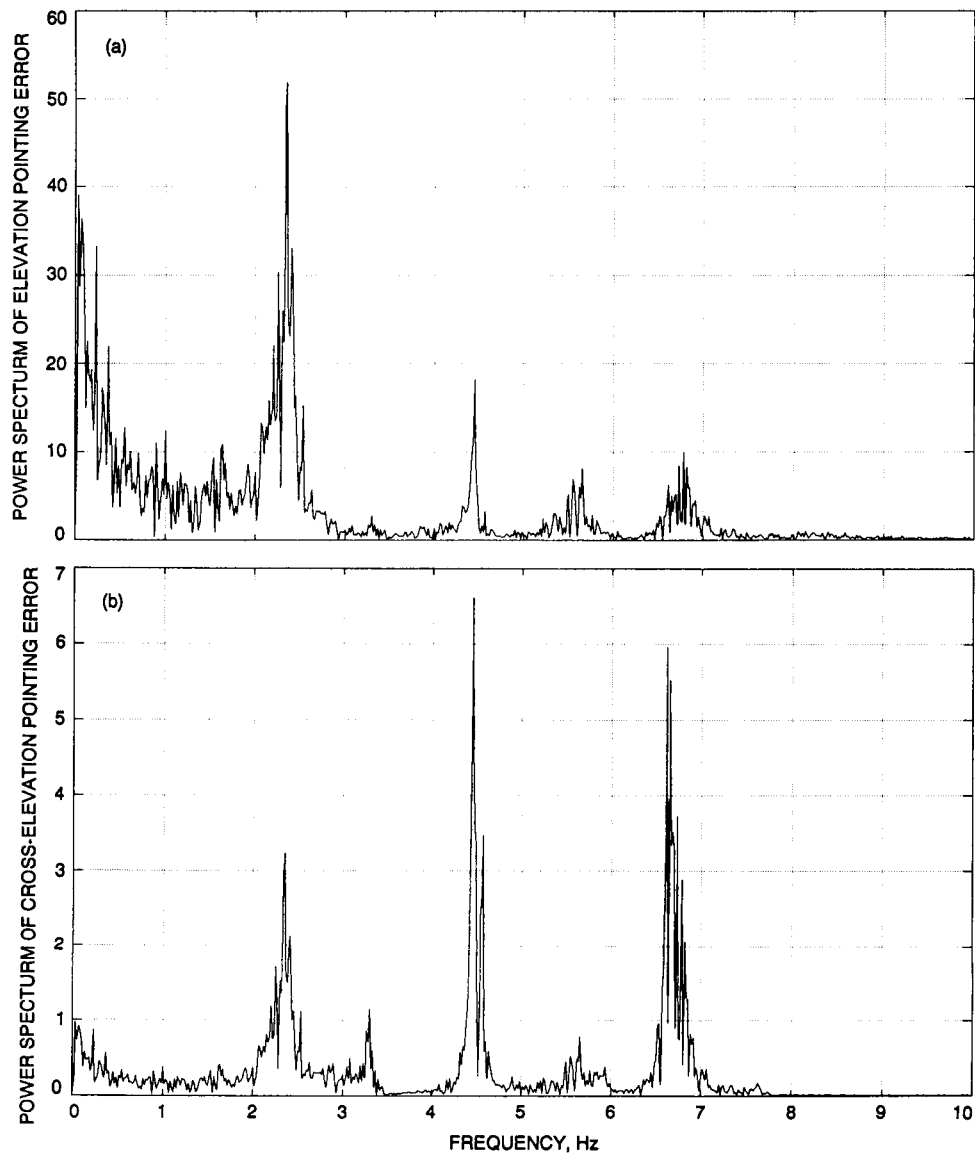


Fig. 10. Spectra for *y*-direction wind gusts: (a) elevation pointing error; (b) cross-elevation pointing error; (c) elevation encoder reading; and (d) azimuth encoder reading.

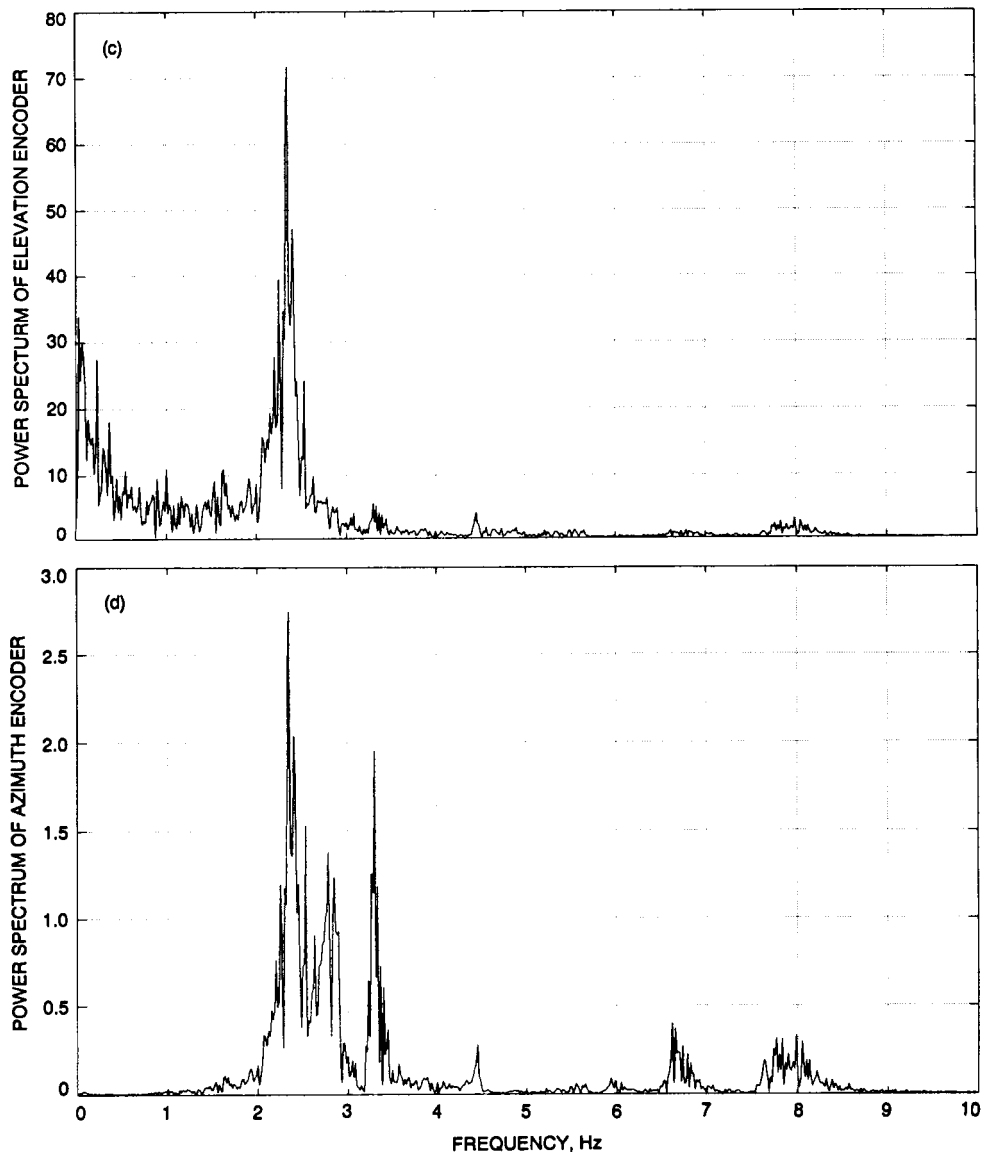


Fig. 10 (contd).

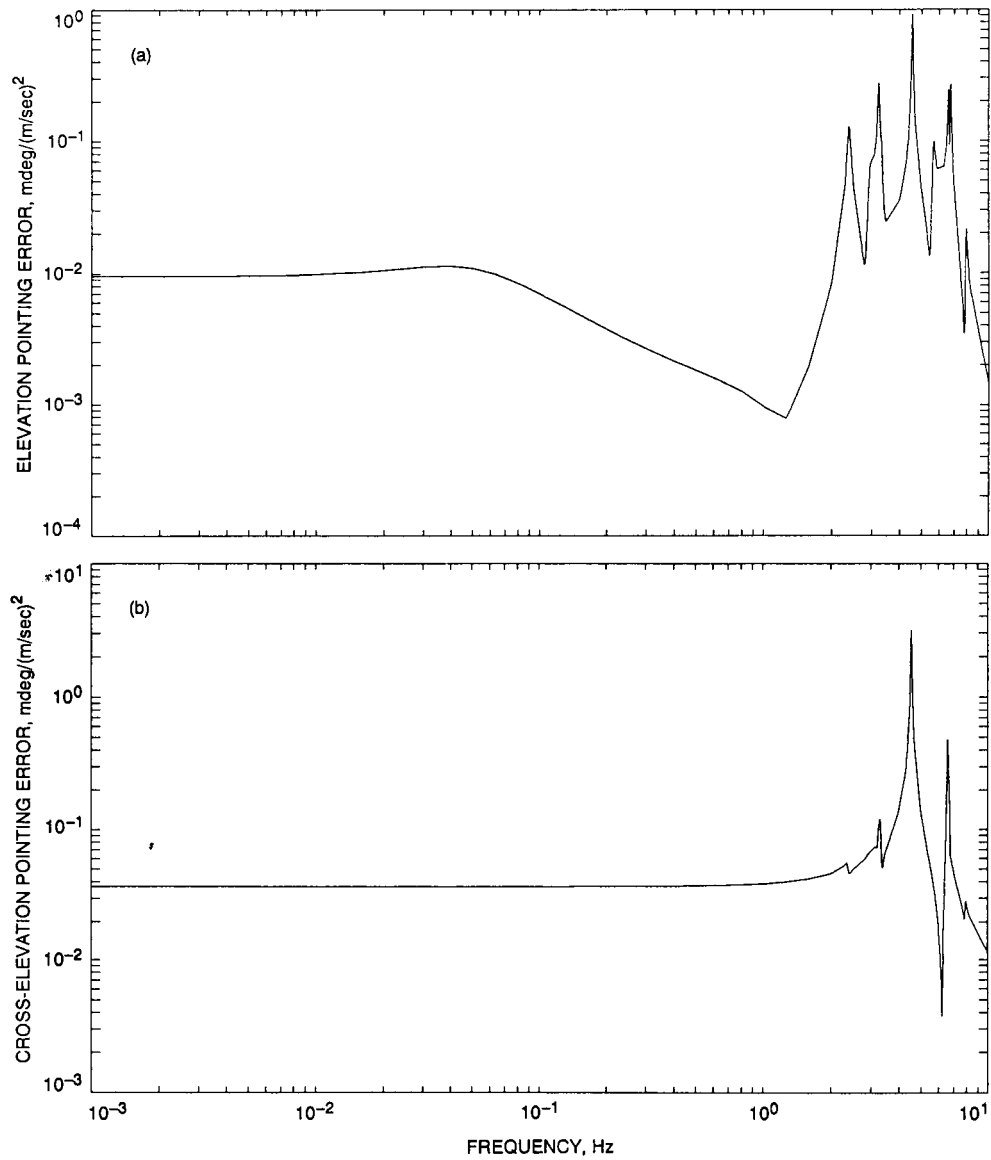


Fig. 11. Frequency-response functions for x-direction wind: (a) elevation pointing error and (b) cross-elevation pointing error.

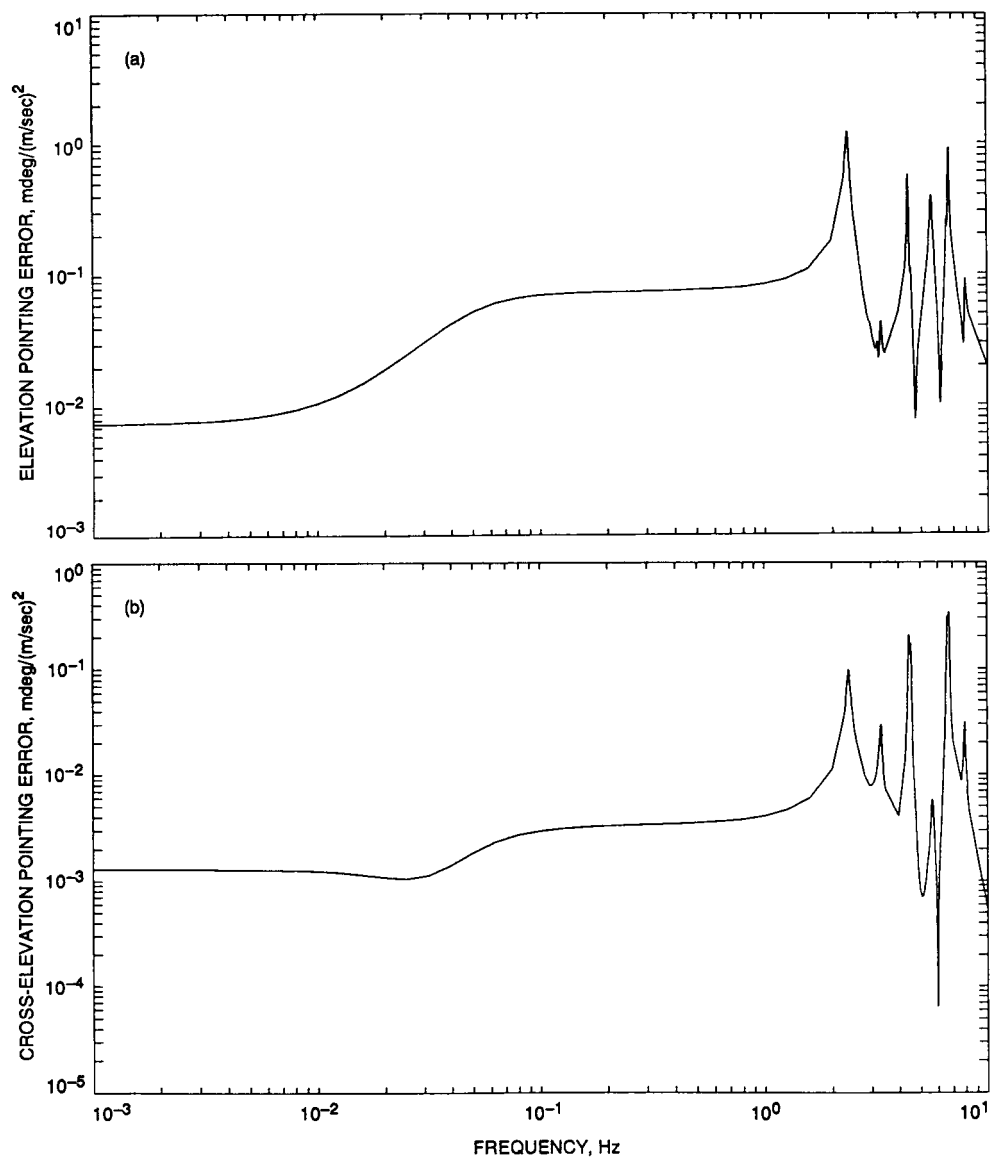


Fig. 12. Frequency-response functions for y-direction wind: (a) elevation pointing error and (b) cross-elevation pointing error.

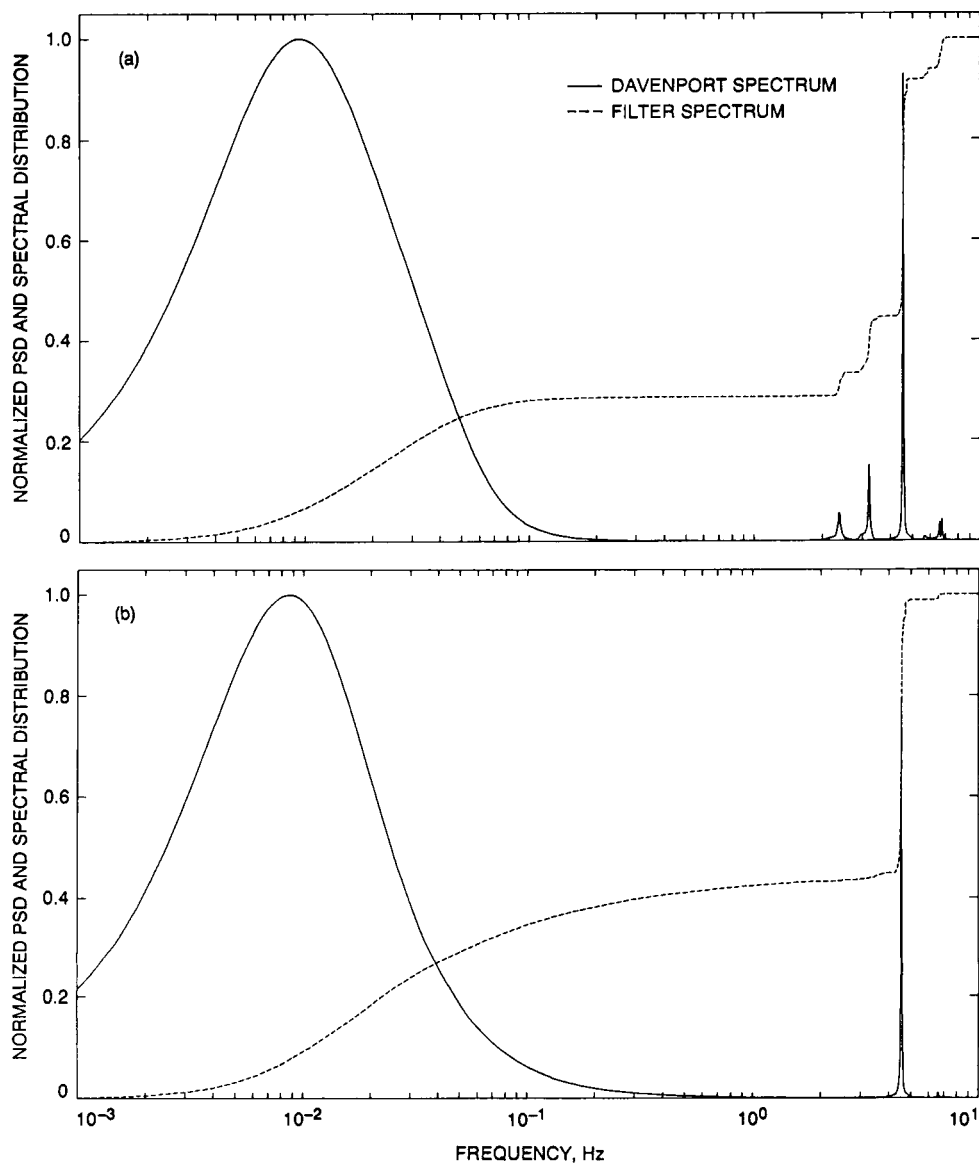


Fig. 13. Spectral density and spectral distribution functions for x -direction wind: (a) elevation pointing error and (b) cross-elevation pointing error.

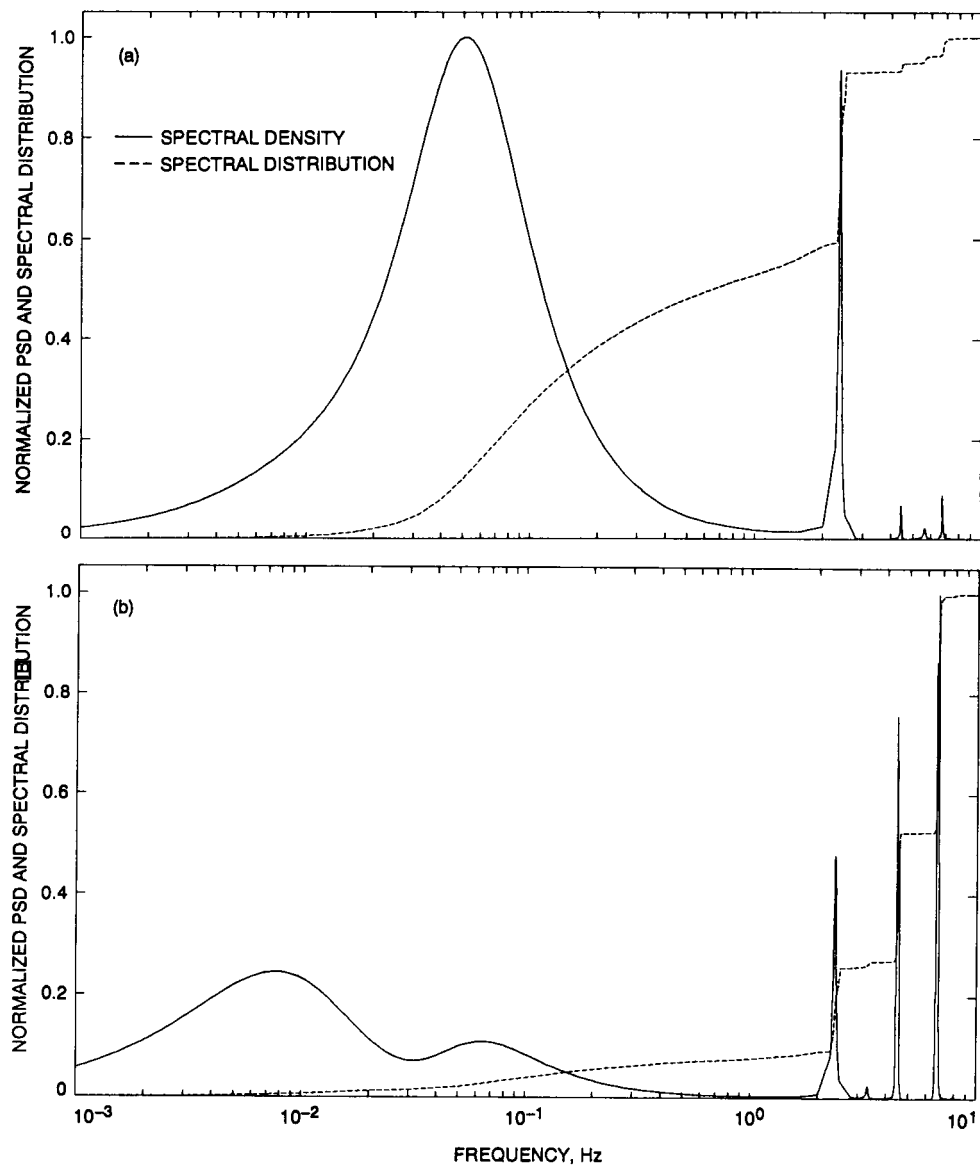


Fig. 14. Spectral density and spectral distribution functions for y-direction wind: (a) elevation pointing error and (b) cross-elevation pointing error.

211-32
77877
P-19
N92-24312
JJ576450

DSS-13 Beam-Waveguide Antenna Performance in the Bypass Mode

S. R. Stewart

Ground Antennas and Facilities Engineering Section and PRC, Inc., Pasadena, California

A new 34-meter beam-waveguide (BWG) antenna that contains two microwave paths, a centerline feed system, and a bypass feed system, has been built at Deep Space Station 13 (DSS 13) at Goldstone, California. Previous articles have described the test results from the evaluation of the centerline BWG feed system in the receive mode as well as the test package hardware used to perform these tests [1-5]. This article presents the test results from the evaluation of the bypass BWG feed system on the DSS-13 antenna in the receive mode, including the operating noise-temperature and the antenna-area-efficiency measurements.

I. Introduction

The new 34-meter beam-waveguide (BWG) antenna, which was built at DSS 13, was designed with two BWG receiver paths: a centerline feed system, which terminates in the pedestal room at the f3 focal point, and a bypass feed system, which terminates on the alidade structure at the f4 focal point (Fig. 1). The centerline feed system has already been evaluated for zenith operating noise-temperature and antenna-area-efficiency performance at X-band (8.45 GHz) and Ka-band (32 GHz) at the two focal points, f1 and f3 [1]. The bypass BWG feed system has now been evaluated at both X-band and Ka-band. In order to maintain consistency, the same test packages that were used to evaluate the centerline feed BWG system were used to test the bypass BWG feed system [2-6].

II. Initial Measurements

Some initial testing at the f4 focal point was conducted to ensure that the receiver equipment was working

properly and that the four mirrors in the bypass system were aligned properly. The X-band test package with its 29-dBi horn was installed at the f4 focal point (Fig. 2) and antenna-efficiency and noise-temperature measurements were immediately begun. Based on previous measurements at f1 and f3 in the centerline-feed BWG mode [1], it was estimated that a well-aligned BWG feed system would result in a peak antenna-aperture efficiency of at least 70 percent near 45 deg elevation. The preliminary data, however, showed a distorted antenna pattern with a peak antenna efficiency of 66 percent at 70 deg elevation. In an effort to find the cause of this problem, the test package was removed and theodolite measurements of the alignment of the mirrors in the bypass feed system were repeated with the antenna at 45 deg elevation. The alignment and positions of the two paraboloid mirrors and the lower flat mirror relative to each other and relative to the shroud walls were verified to be correct to within ± 0.40 cm from their design positions and ± 0.02 deg from their design angles with the antenna at 45 deg elevation. The position and angle of the upper flat mirror and the relative position

between the upper paraboloid and flat mirrors were verified with the antenna at zenith. The measurements that could not be made, however, were the changes in relative position between the mirrors due to flexure in the antenna structure as the antenna moves in elevation.

Knowing that all of the hardware was positioned correctly, a detailed investigation of the antenna's performance was begun by measuring its far-field pattern. The high signal-to-noise ratio of the 12.198-GHz beacon on board a geostationary satellite was required for accurate antenna-pattern measurements, so a Ku-band (12.198-GHz) receiver system [6] was installed at f4. Antenna-pattern measurements were made that indicated a substantial misalignment in the mirrors relative to their optimum positions (Fig. 3).

A full analysis of the alignment errors contributed by each of the four mirrors would have seriously delayed this work, so it was decided to move the most sensitive mirror until the antenna pattern, with the antenna near 45 deg elevation, was balanced and then proceed with the antenna-efficiency measurements even though it was known that movement of just one mirror would not yield the optimum operating noise temperature and antenna gain.

Communication with D. Bathker and W. Veruttipong¹ indicated that the paraboloid mirror on the dish surface was the most sensitive. Movements of both 0.193 deg about the mirror's x-axis and 0.255 deg about the y-axis were required to balance the antenna pattern (Figs. 4 and 5). The cause for this discrepancy between the optical alignment and the RF alignment is presumed to be at least partially due to flexure in the antenna between the 45- and 90-deg elevation angles, where the mirror alignments were checked. After the alignment was complete, the X-band test package was reinstalled at f4 and operating noise-temperature and antenna-efficiency measurements resumed.

III. Measurement Techniques

Throughout this experiment, two radiometer programs were relied on for maximizing the antenna performance measurements. The software is almost identical to that which was used during the testing of the centerline BWG system, with a few minor modifications and improvements. The first software package is the automatic bore-sight program, AUTOBORE, which measures the radio-source noise temperature and calculates antenna pointing

errors by making a continuous series of five-point bore-sights on a given source throughout a 12-hour track. This software has been described in detail elsewhere [1,4], and some parts are reproduced below.

The second package is the calibration software, which is used to measure the gain stability, amplifier linearity, and operating temperature of the system over long periods of time by making calibration measurements periodically and recording these data to disk. The receiver is calibrated by measuring the detected noise power with the receiver in four different configurations: the receiving path of the waveguide switch turned to the feed horn, the receiving path turned to a waveguide ambient load, and each of these configurations with an injected noise-diode signal. A power measurement with the power meter terminated into a 50- Ω load and a physical temperature measurement of the ambient load are also made and added to a disk file. Calculations of the system gain based on the known temperature and noise-power output of the waveguide ambient load are made and used to determine the current operating noise temperature. Evaluation of system linearity is made by comparing the noise power injected by the diode in each of the waveguide configurations. These calibrations are made automatically every 30 minutes or at any other user-defined time interval.

IV. Test Results

A. X-Band Zenith Noise-Temperature Measurements

The X-band test package with its 29-dBi horn was used to make comparative zenith operating noise-temperature measurements at two locations: on the ground and on the antenna at the f4 focal point. Continuous measurements at each of these locations were conducted over two-day periods, with operating noise-temperature (T_{op}) measurements being made every 30 minutes and ambient weather conditions, including temperature, pressure, and relative humidity (RH), being measured every 10 minutes. Figures 6(a) and (b) contain typical calibration data that were collected with the X-band test package on the ground, and Figs. 7(a) and (b) contain calibration data collected at f4, with the antenna at the zenith position. These calibration plots are included to exemplify the stability of the system during the measurement period. Since each of the operating noise temperatures was calculated using the most recent gain factor, corrections for gain changes were automatically accounted for. Corrections for system nonlinearity, however, were not made since the magnitude of the linearity factor is on the order of its uncertainty at ± 1 percent and making any corrections for this factor would have only a minor effect. Table 1 shows the compiled results

¹ Private communications with D. Bathker and W. Veruttipong of the Jet Propulsion Laboratory on 11 April 1991 and 19 April 1991.

would have only a minor effect. Table 1 shows the compiled results of the X-band T_{op} measurements, as well as the weather conditions present during the observation period. Table 1 also shows the corrections that were made to the operating noise-temperature measurements in order to normalize them to the weather conditions present during an average Goldstone day. This normalized T_{op} was computed from

$$T_{op,n} = T_{op} + (T'_{cb}/L_{wg})(1/L_{atm,s} - 1/L_{atm}) \\ + (1/L_{wg})(T_{atm,s} - T_{atm}) + (T_{wg,s} - T_{wg}) \quad (1)$$

where T_{atm} and T_{wg} are, respectively, the atmospheric and waveguide noise-temperature contributions, and L_{atm} and L_{wg} are the atmospheric and waveguide loss factors.

Values for the standard DSS-13 atmospheric conditions² at 8.45 GHz are

$$T_{atm,s} = 2.17 \text{ K}$$

$$L_{atm,s} = 1.00814 \text{ (corresponding to 0.0352 dB)}$$

Other X-band values used for the above equation are

$$T'_{cb}{}^3 = 2.5 \text{ K}$$

$$L_{wg} = 1.0163 \text{ (corresponding to 0.07 dB)}$$

$$T_{wg,s} = 4.69 \text{ K for the above } L_{wg} \text{ and a standard physical waveguide temperature of } 20 \text{ deg C}$$

By normalizing the T_{op} measurements, it is easier to compare these measurements with each other as well as with those made at any other time of the year. As a means of comparison between these results for f4 and the results obtained previously for f1 and f3, Table 2 contains a summary of some differential zenith operating noise temperatures. All of the T_{op} values presented in Table 2 have been normalized to standard DSS 13 atmospheric conditions at 8.45 GHz.

B. X-Band Antenna-Efficiency Measurements

The antenna's efficiency performance is determined by continuously tracking a stellar radio source and measur-

ing the peak received noise power in relation to the background atmospheric noise over a full 10–12 hour track. By knowing the maximum signal strength of the radio source, the aperture efficiency with respect to the elevation angle of the antenna can be determined.

During X-band testing, eight full days of antenna-efficiency measurements were conducted, using the four radio sources that are listed and described in Table 3. Two of these sources are constant and well calibrated while the other two are variable. The two variable sources are good for relative measurements over a short period of days, but need to be calibrated against known, stable sources. Therefore, Table 4 contains the correction factors which needed to be applied to the variable sources in order to make them useful. Figure 8 shows the combined antenna efficiency measurements at X-band versus elevation angle for the f4 focal point with the effects of atmospheric attenuation removed. An average antenna efficiency with a peak of 71.4 percent at an elevation angle of 38.6 deg is also shown in this figure.

C. Ka-Band Zenith Noise-Temperature Measurements

As with X-band, the Ka-band test package with its 29-dBi horn was used to make comparative zenith T_{op} measurements on the ground and on the antenna at the f4 focal point (Fig. 9). Continuous measurements at each of these locations were made over two-day periods, with operating noise-temperature measurements being taken every 30 minutes and ambient weather conditions being measured every ten minutes. Figures 10(a) and (b) contain typical calibration data that were collected with the Ka-band test package on the ground, and Figs. 11(a) and (b) contain typical calibration data collected at f4, with the antenna at the zenith position. As with the X-band measurements described above, corrections for gain changes are already accounted for in the operating noise-temperature measurements, and no corrections were made for linearity which would have an effect of less than ± 2 percent.

Table 1 shows the compiled results of the Ka-band T_{op} measurements, as well as the weather conditions present during the observation period. The normalized T_{op} values, which were calculated using Eq. (1), are also included in this table, for comparison purposes.

Values for the standard DSS-13 atmospheric conditions at 32 GHz are

$$T_{atm,s} = 7.02 \text{ K}$$

$$L_{atm,s} = 1.02683 \text{ (corresponding to 0.1150 dB)}$$

² Deep Space Network/Flight Project Interface Design Handbook, Document 810-5, Rev. D, vol. I, sec. TCI-30 (internal document), Jet Propulsion Laboratory, Pasadena, California, June 1, 1990.

³ T'_{cb} is the effective noise contribution from cosmic background radiation. This value is a function of frequency and will differ from the nominal noise temperature of 2.7 K.

Other Ka-band values used for Eq. (1) are

$$T'_{cb} = 2.0 \text{ K}$$

$$L_{wg} = 1.06414 \text{ (corresponding to 0.27 dB)}$$

$$T_{wg,s} = 17.67 \text{ K for the above } L_{wg} \text{ and a standard physical waveguide temperature of } 20 \text{ deg C}$$

Table 2 contains a comparison between these results for f4 and the results obtained previously for f1 and f3. This comparison may be slightly erroneous due to the six-month time lapse between measurements. All of the T_{op} values presented in Table 2 have been normalized to standard DSS-13 atmospheric conditions at 32 GHz.

D. Ka-Band Antenna-Efficiency Measurements

During Ka-band testing, eight full days of antenna-efficiency measurements were conducted using four radio sources. These four sources are listed and described in Table 5. The flux densities of two sources, 3C 274 and Venus, have been thoroughly measured at 32 GHz, but those of the other two sources have not. In addition, flux values from Venus are continually changing due to the change in Venus's position relative to the Earth, and must be updated daily. The two variable sources that were used (3C 84 and 3C 273) had to have their flux

values corrected against the calibrated sources, and these corrections are found in Table 6. Figure 12 shows the combined antenna-efficiency measurements versus elevation angle for Ka-band at the f4 focal point, with the effects of atmospheric attenuation removed. An average antenna-efficiency having a peak of 45.0 percent at an elevation angle of 41.4 deg is also shown in this figure. The standard deviation of these data points is 5.0 percent.

Figure 13 contains the combined results of antenna-efficiency measurements at both X-band and Ka-band for the three focal points, f1, f3, and f4. The results shown in this figure, like those in Figs. 8 and 12, have had the effects of atmospheric attenuation removed.

V. Conclusions

The bypass feed system of the DSS-13 BWG antenna has been fully tested for noise-temperature and antenna-efficiency performance at both X-band and Ka-band. The mirror movements inside the beam waveguide that were necessary in order to make adequate measurements of these parameters resulted in a bypass feed system which may be slightly less than optimal. The results, however, show a system whose performance at both 8.45 and 32 GHz at the bypass focal point, f4, is consistent with those at f1 and f3.

Acknowledgments

The assistance of DSS-13 personnel G. Bury, J. Crook, L. Smith, R. Reese, J. Garnica, G. Farner, and R. Littlefair is gratefully acknowledged. The author also thanks DSS-13 group supervisor, C. Goodson, and the DSS-13 station manager, A. Price, for their continual support throughout the measurement process.

References

- [1] S. D. Slobin, T. Y. Otoshi, M. J. Britcliffe, L. S. Alvarez, S. R. Stewart, and M. M. Franco, "Efficiency Calibration of the DSS 13 34-Meter Diameter Beam Waveguide Antenna at 8.45 and 32 GHz," *TDA Progress Report 42-106*, vol. April-June 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 283-297, August 15, 1991.
- [2] T. Y. Otoshi, S. R. Stewart, and M. M. Franco, "A Portable X-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation—Part I: Design and Ground Tests," *TDA Progress Report 42-103*, vol. July-September 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 135-150, November 15, 1990.
- [3] T. Y. Otoshi, S. R. Stewart, and M. M. Franco, "A Portable X-Band Front-End Test Package for Beam Waveguide Antenna Performance Evaluation—Part II: Tests on the Antenna," *TDA Progress Report 42-105*, vol. January-March 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 54-68, May 15, 1991.
- [4] T. Y. Otoshi, S. R. Stewart, and M. M. Franco, "A Portable Ka-Band Front-End Test Package for Beam Waveguide Antenna Performance Evaluation—Part I: Design and Ground Tests," *TDA Progress Report 42-106*, vol. April-June 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 249-265, August 15, 1991.
- [5] T. Y. Otoshi, S. R. Stewart, and M. M. Franco, "A Portable Ka-Band Front-End Test Package for Beam Waveguide Antenna Performance Evaluation—Part II: Tests on the Antenna," *TDA Progress Report 42-106*, vol. April-June 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 266-282, August 15, 1991.
- [6] T. Y. Otoshi, S. R. Stewart, and M. M. Franco, "A Portable Ku-Band Front-End Test Package for Beam Waveguide Antenna Performance Evaluation," *TDA Progress Report 42-107*, vol. July-September 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 73-80, November 15, 1991.

Table 1. Measured zenith operating noise temperatures corrected for weather and waveguide-loss changes for X-band and Ka-band.

Configuration	Observation period, UT	Average measured T_{op} , K	Average weather during observation	Computed T_{atm} , K	Computed L_{atm}	Physical waveguide temp., deg C	T_{wg} , K	Normalized T_{op} , K
X-band, on the ground	03/12/91 1600 03/13/91 1730	22.29	893.6 mbar 8.9 deg C 37.3% RH	2.25	1.0085 (0.037 dB)	13.34	4.58	22.32
X-band, at f4	04/27/91 0500 04/28/91 2030	28.84	891.3 mbar 17.6 deg C 26.4% RH	2.27	1.0086 (0.037 dB)	13.96	4.59	28.84
Ka-band, on the ground	04/23/91 1200 04/24/91 2330	83.09	890.8 mbar 9.96 deg C 69.9% RH	10.94	1.0418 (0.178 dB)	5.05	16.77	80.33
Ka-band, at f4	05/11/91 0030 05/12/91 2230	93.51	894.1 mbar 12.6 deg C 34.4% RH	8.22	1.0312 (0.133 dB)	7.65	16.93	93.13

Table 2. Differential zenith operating noise temperatures at X-band.

Configurations differenced	X-band delta T_{op} , K	Ka-band delta T_{op} , K
F1-Ground	3.2 ^a	7.1 ^b
F2-F1	—	5.2 ^b
F3-F1	8.9 ^a	6.8 ^b
F4-F1	3.3	5.7
F2-Ground	—	12.3
F3-Ground	12.1	13.9
F4-Ground	6.5	12.8

^a These values obtained from [3].
^b These values obtained from [5].

Table 3. Radio sources used for X-band calibrations at f4, April 1991. (The values below were obtained from [1].)

Source	Declination, J2000.0	Peak elevation at DSS 13	Flux, Jy	C_r	T_{100}/C_r , K
3C 274	12.391	67.1	44.555	1.087	13.477
3C 123	29.671	84.4	9.404	1.0054	3.075
3C 84	41.512	83.7	45.79	1.000	15.056
3C 273	2.052	56.8	36.46	1.000	11.988

Table 4. Efficiency adjustments, deduced flux, and T100/C_r values for X-band at f4, April 1991.

Source	Efficiency adjustment needed	Deduced flux, Jy	Deduced T100/C _r , K
3C 274 ^a	1.000	44.555	13.477
3C 123 ^b	0.98558	9.542	3.12
3C 84 ^c	1.43801	31.843	10.47
3C 273 ^c	1.17529	31.022	10.20

^a Constant, extended.

^b Constant, nonpoint.

^c Variable, point.

Table 5. Radio sources used for Ka-band calibrations at f4, May 1991.

Source	Declination, J2000.0	Distance, AU	Peak elevation at DSS 13	Flux, Jy	C _r	T100/C _r , K
3C 274 ^a	12.391		67.1	16.22	1.2730	4.190
3C 84 ^a	41.512		83.7	43.71	1.0000	14.372
3C 273 ^a	2.052		56.8	29.54	1.0000	9.713
Venus 1991 ^b						
01 May	+25.3	1.0511	80.0	71.085	1.0254	22.804
10 May	+25.9	0.9842	80.6	81.076	1.0290	25.918
20 May	+25.5	0.9072	80.2	95.430	1.0342	30.353

^a Values obtained from [1].

^b Declination of date.

Table 6. Efficiency adjustments, deduced flux, and T100/C_r values for Ka-band at f4, May 1991.

Source	Efficiency adjustment needed	Deduced flux, Jy	Deduced T100/C _r , K
3C 274 ^a	0.98172	16.52	4.268
Venus ^b	1.0000	See Table 5	See Table 5
3C 84 ^c	1.45687	30.003	9.865
3C 273 ^c	0.62867	46.987	15.45

^a Constant, extended.

^b Ephemeris dependent, nonpoint.

^c Variable, point.

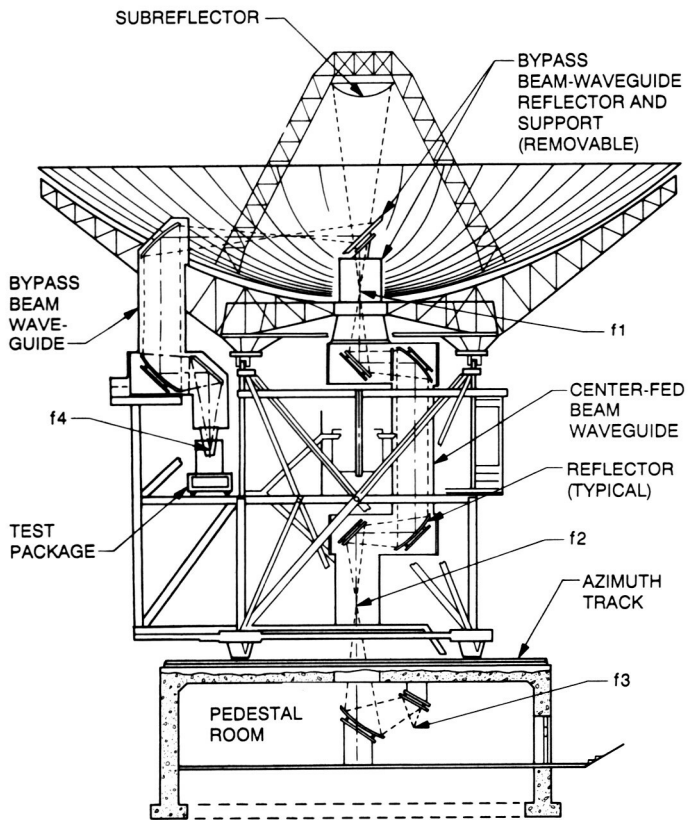


Fig. 1. DSS-13 beam-waveguide antenna depicting focal points f_1 , f_3 , and f_4 .



Fig. 2. X-band test package with its 29-dBi horn being lifted into position at the f_4 focal point.

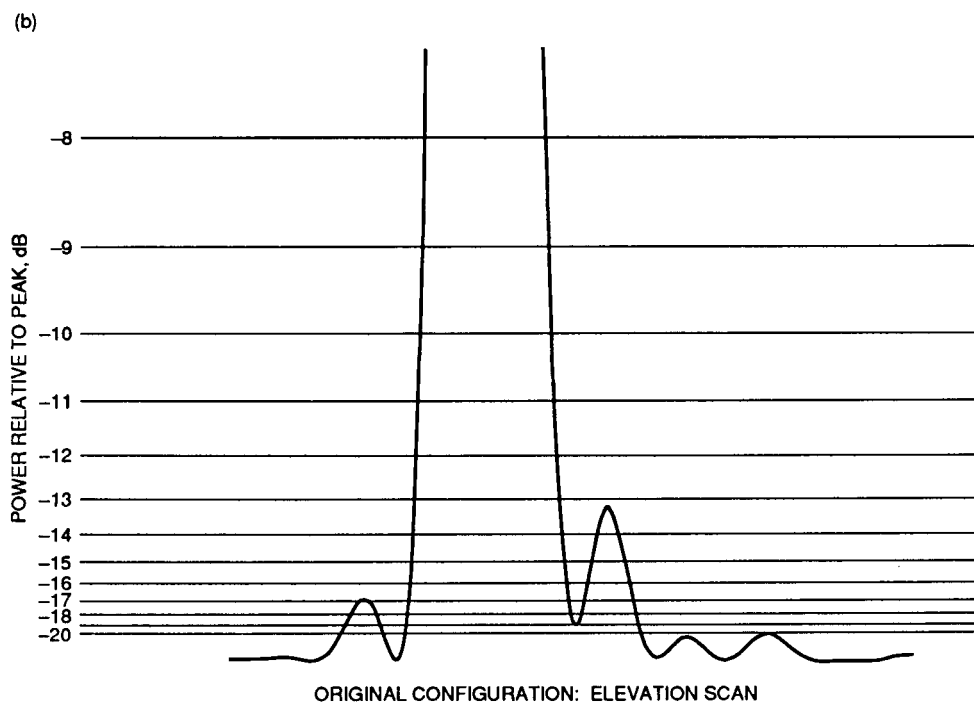
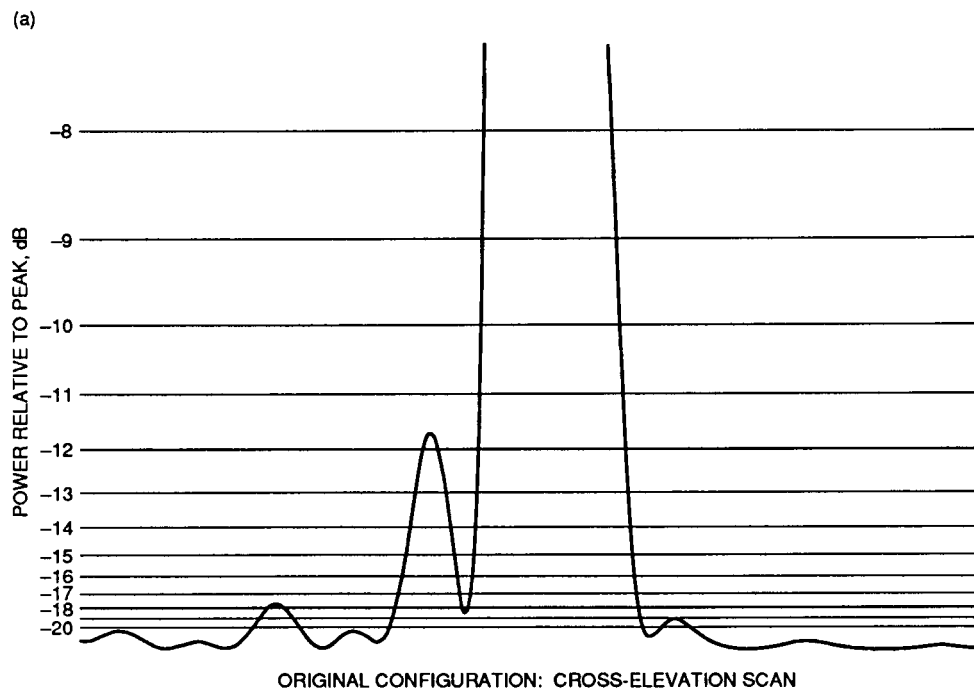


Fig. 3. Antenna pattern before mirror alignment: (a) in the cross-elevation axis and (b) in the elevation axis.

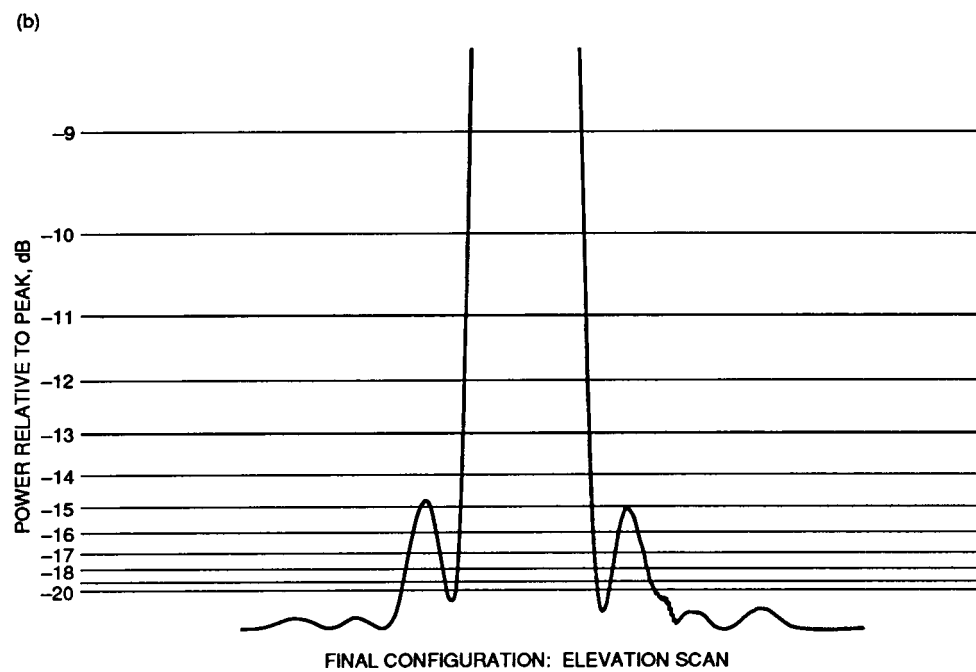
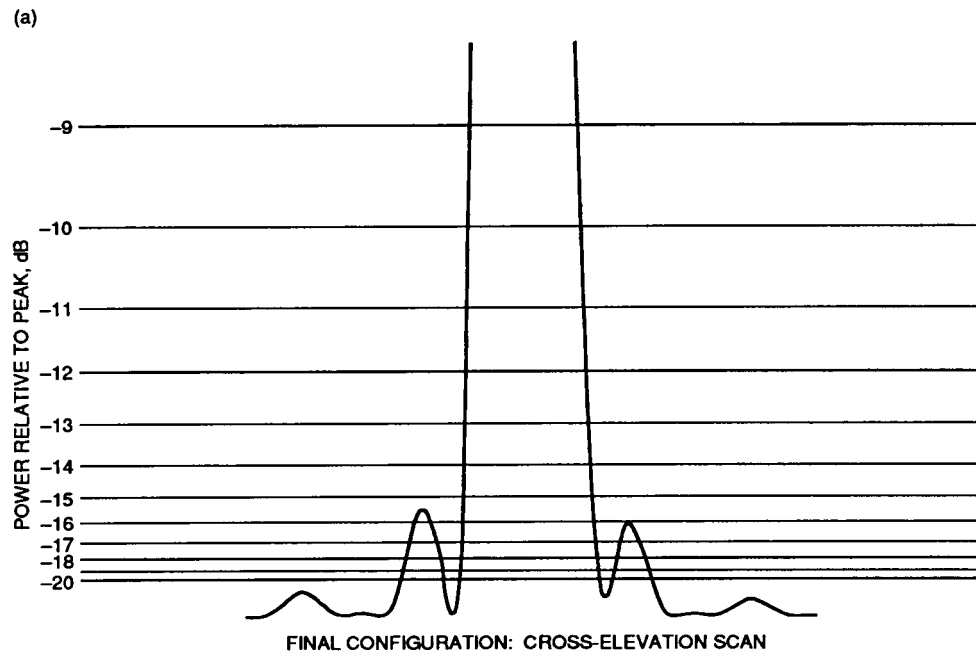


Fig. 4. Antenna pattern after mirror alignment: (a) in the cross-elevation axis and (b) in the elevation axis.

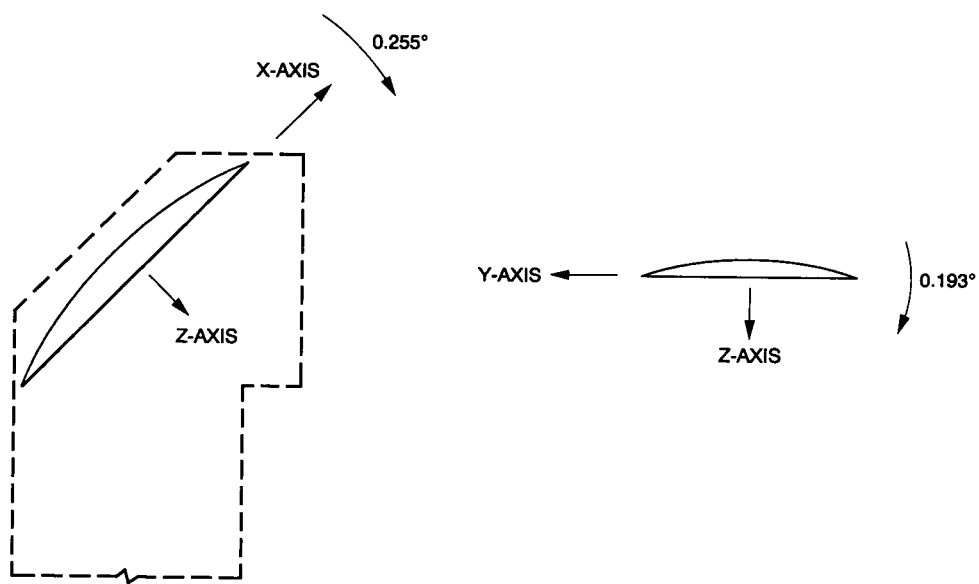


Fig. 5. Motion of upper parabolic mirror.

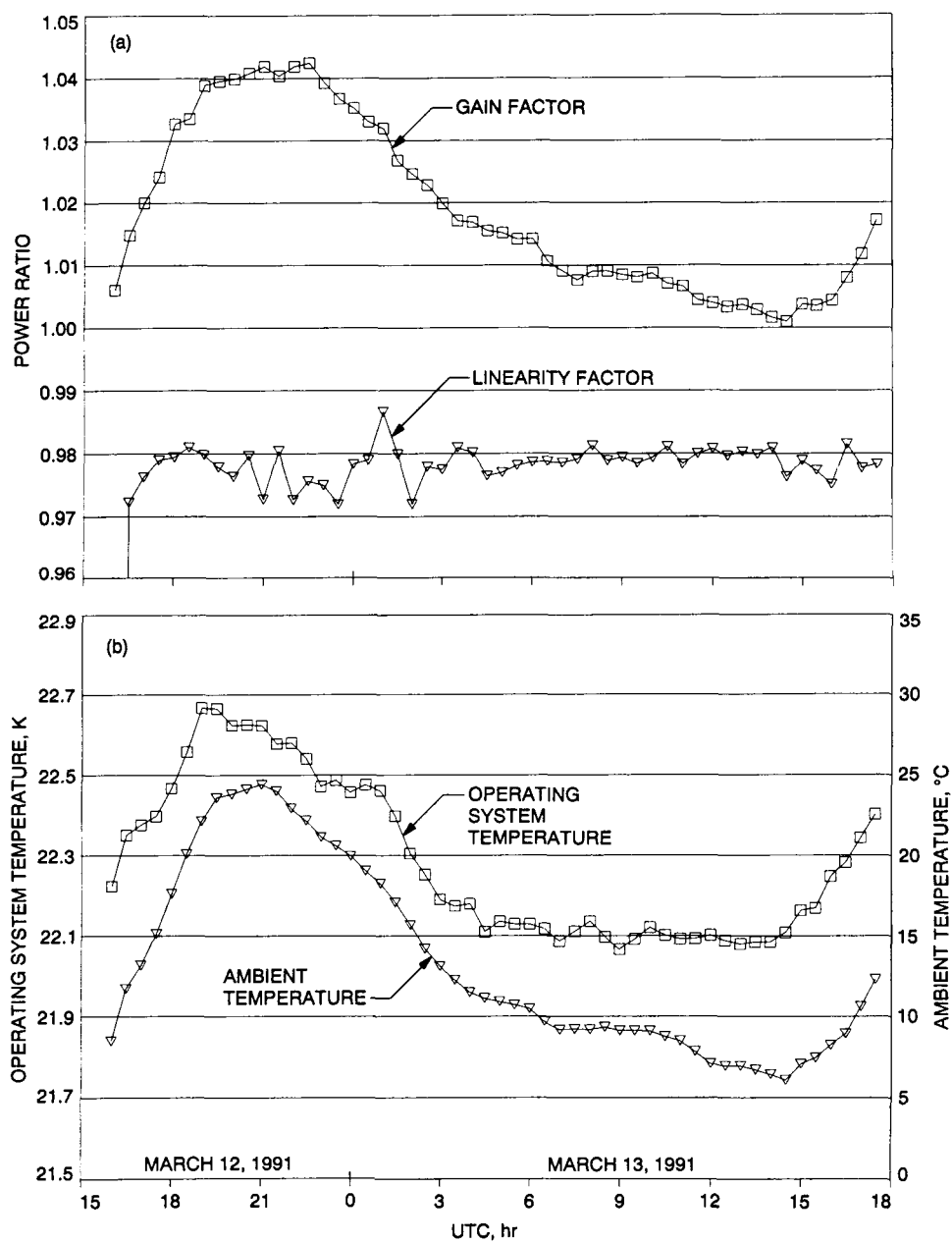


Fig. 6. X-band zenith noise-temperature measurements on the ground: (a) system gain and linearity and (b) system temperature.

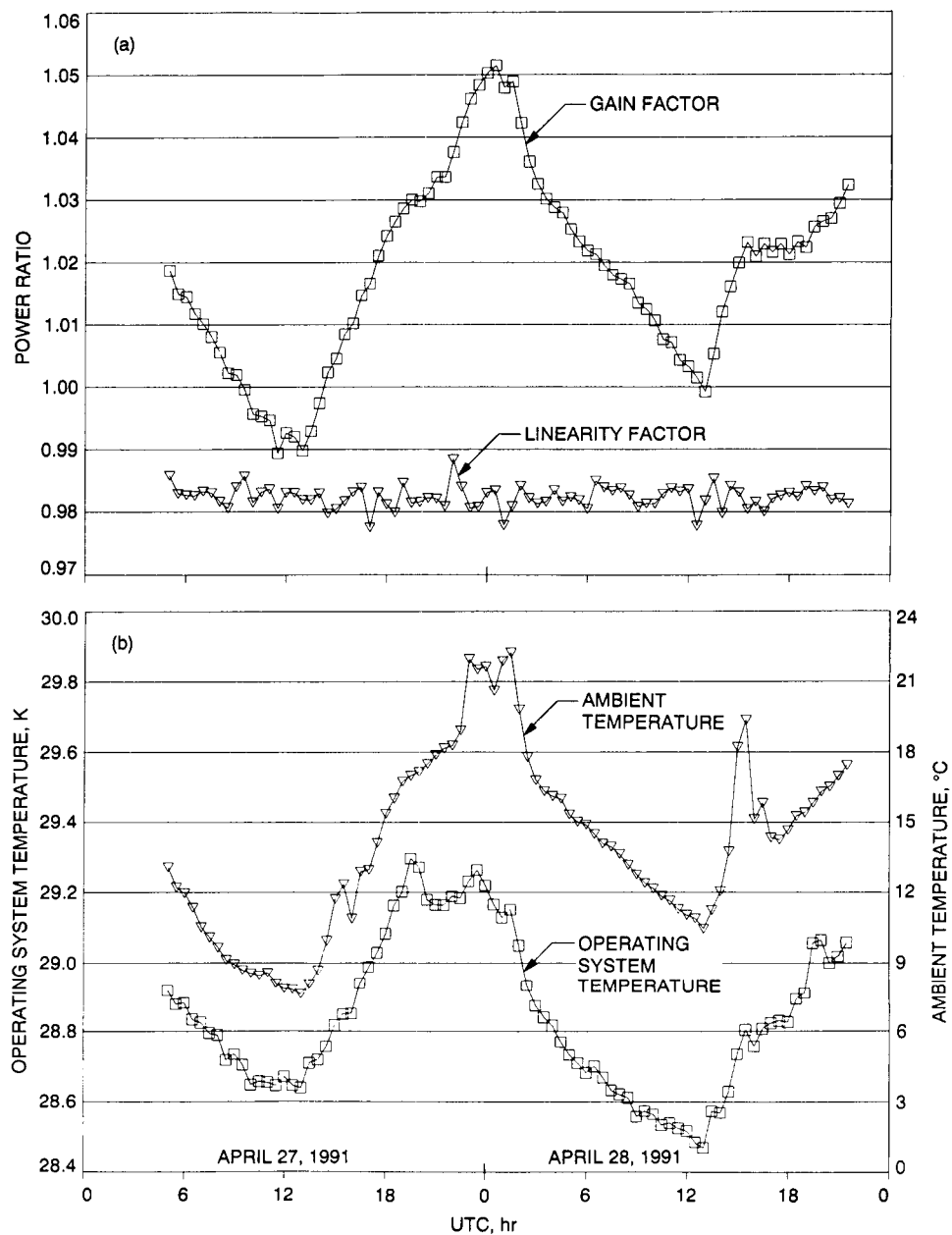


Fig. 7. X-band zenith noise-temperature measurements on the antenna: (a) system gain and linearity and (b) system temperature.

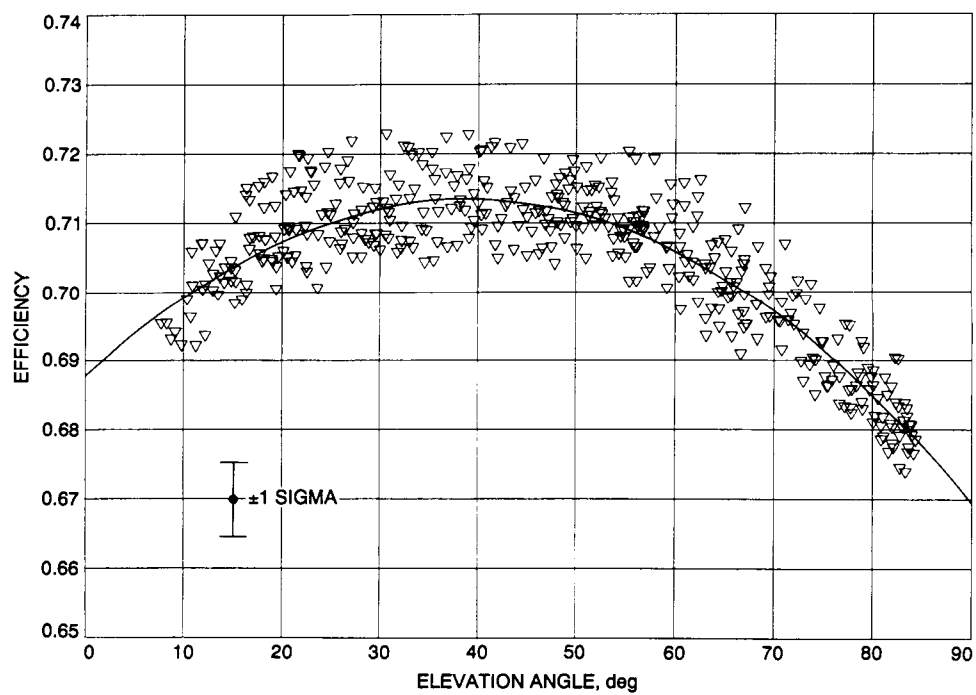


Fig. 8. X-band efficiency plot.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

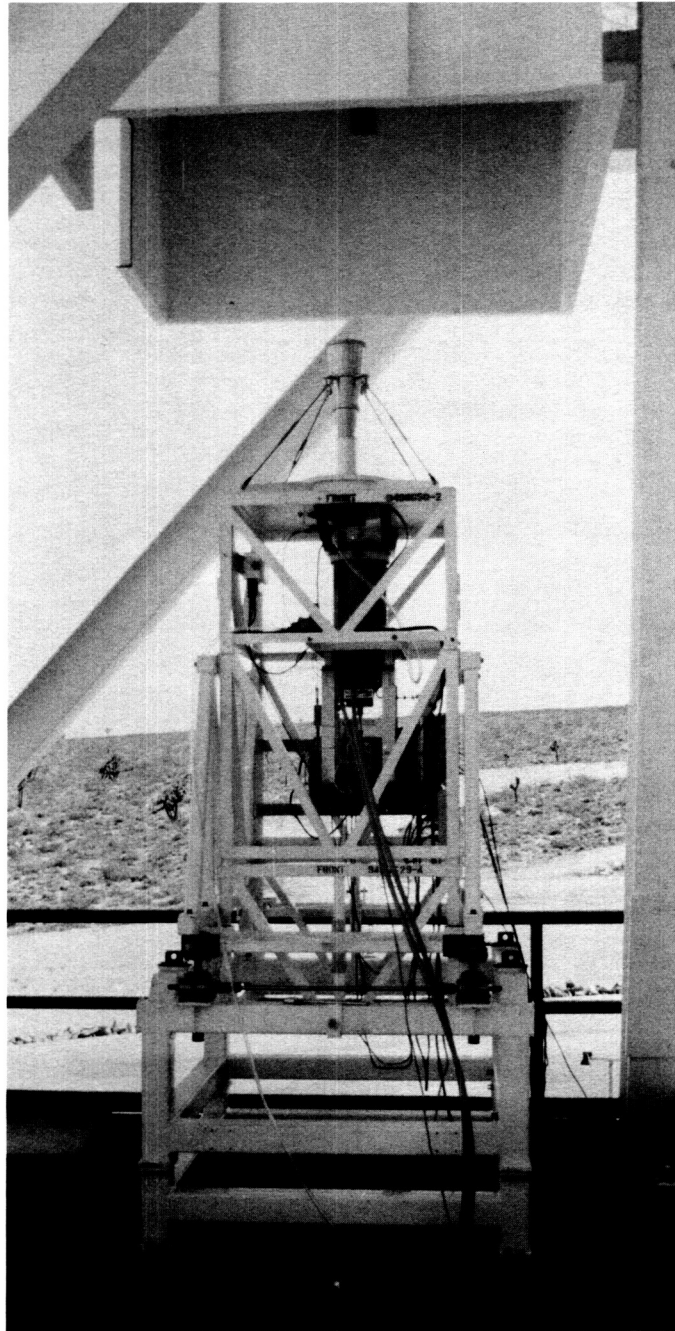


Fig. 9. Ka-band test package with 29-dBi horn at the f_4 focal point.

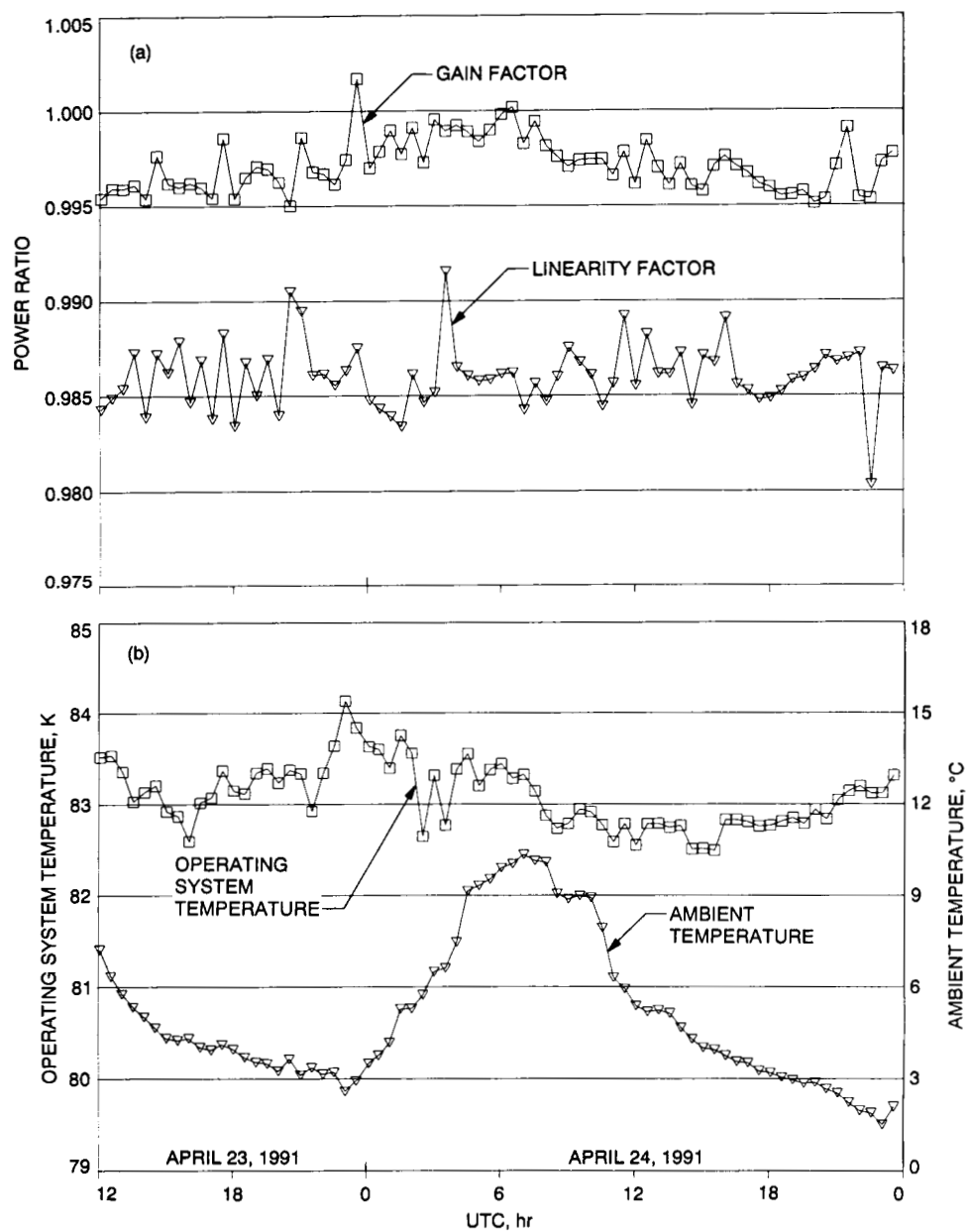


Fig. 10. Ka-band zenith noise-temperature measurements on the ground: (a) system gain and linearity and (b) system temperature.

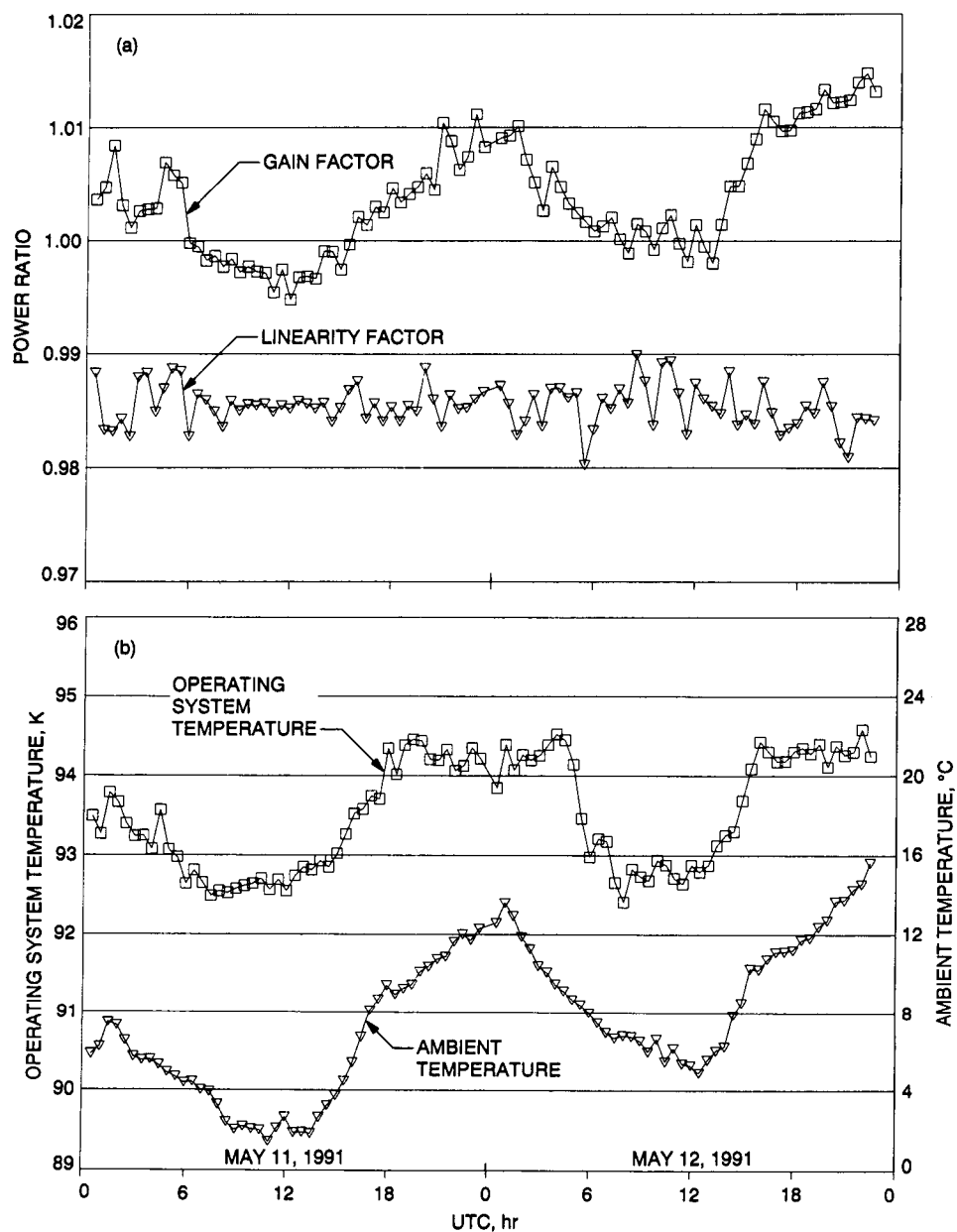


Fig. 11. Ka-band zenith noise-temperature measurements on the antenna: (a) system gain and linearity and (b) system temperature.

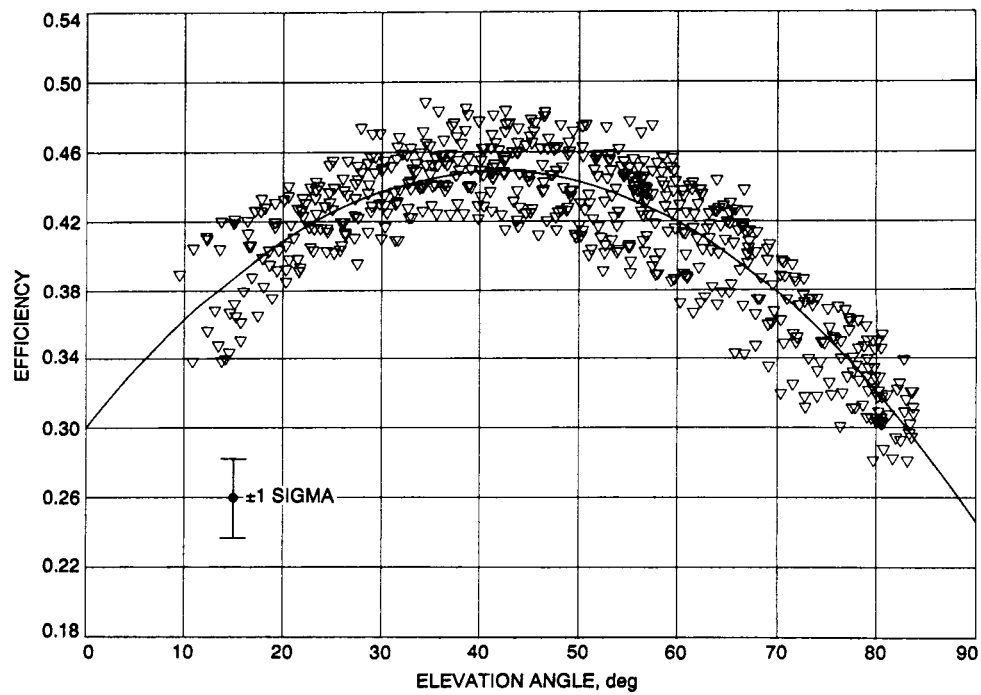


Fig. 12. Plot of Ka-band efficiency.

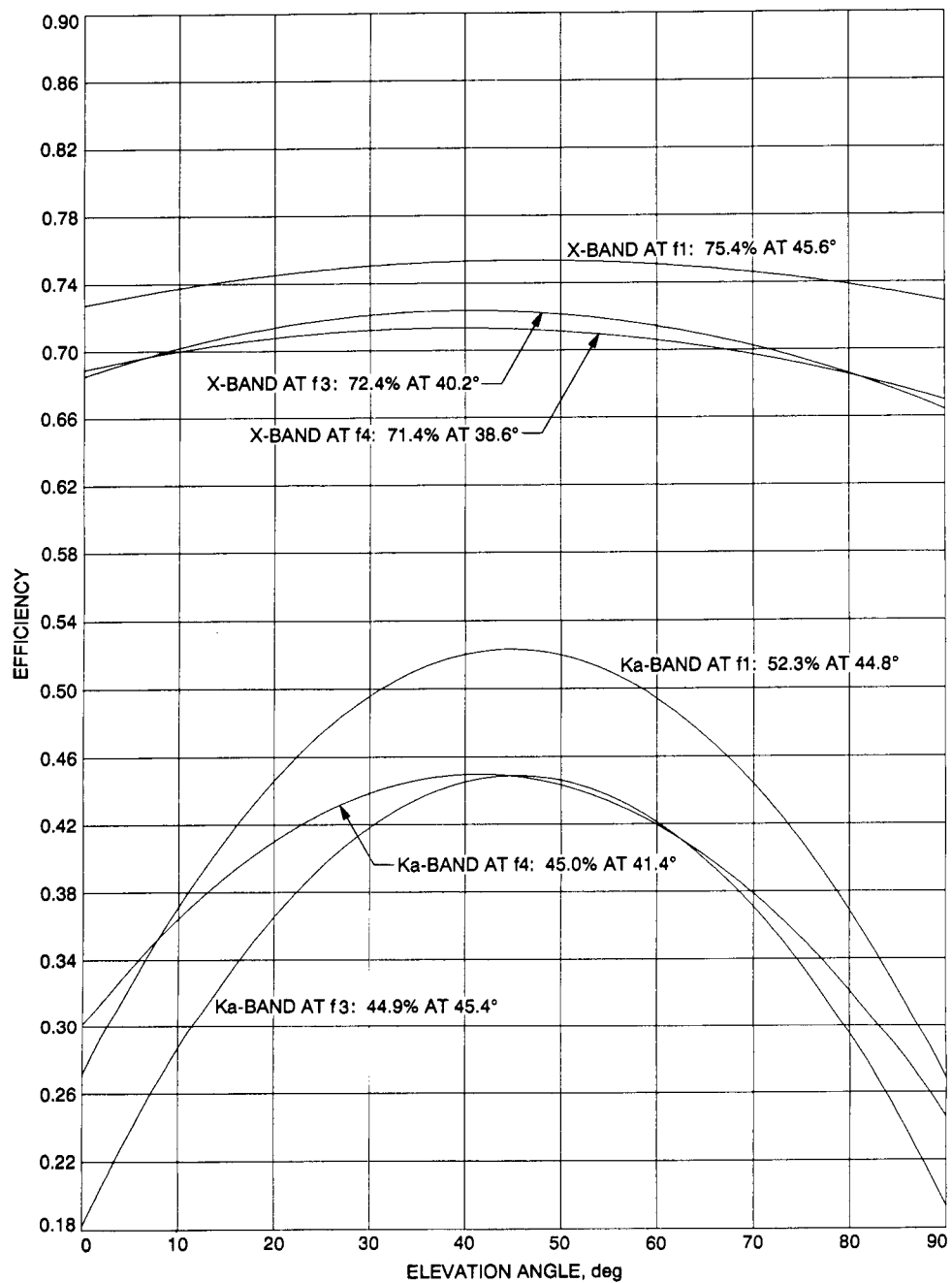


Fig. 13. Comparative results for all of the efficiency measurements on the DSS-13 antenna.

512-32
75872
P-11 N92-24313
JJ576450

The Electrical Conductivities of the DSS-13 Beam-Waveguide Antenna Shroud Material and Other Antenna Reflector Surface Materials

T. Y. Otoshi, M. M. Franco, and H. F. Reilly, Jr.
Ground Antennas and Facilities Engineering Section

A significant amount of noise temperature can potentially be generated by currently unknown dissipative losses in the beam-waveguide (BWG) shroud. The amount of noise-temperature contribution from this source is currently being investigated. In conjunction with this study, electrical conductivity measurements were made on samples of the DSS-13 BWG shroud material at 8.420 GHz. The effective conductivities of unpainted and painted samples of the BWG shroud have been measured to be 0.01×10^7 and 0.0036×10^7 mhos/m, respectively. This value may be compared with 5.66×10^7 mhos/m for high-conductivity copper.

I. Introduction

A description of the DSS-13 beam-waveguide (BWG) antenna as well as experimental noise temperature data have been presented in a Phase I Final Report.¹ Although antenna efficiencies agreed well with predictions, measured operating noise temperatures for the BWG portion of the system at 8.45 GHz were found to be about 6 K higher than had been expected.

The causes of the noise temperatures generated within a BWG antenna are currently under investigation. A possible cause of noise increase being studied is the dissipative

loss in the walls of the BWG steel shroud. The shroud, 2.44 m in diameter and about 24.4 m long, begins at the Cassegrain vertex and, after four 90-deg turns, ends at the antenna azimuth plane. Then, for about 1.5 m, between the azimuth plane down to the ceiling of the pedestal room, the shroud wall is made of concrete instead of steel. The remainder of the microwave path to the final focal point F3 is unshrouded.

Little information is available in the technical literature concerning the electrical conductivity values of various types of steel at microwave frequencies. It is the purpose of this article to summarize all known published data on the conductivity values of many types of steel materials. Most of the values to be presented were obtained from

¹ M. Britcliffe, Editor, *DSS-13 Beam Waveguide Antenna Project: Phase I Final Report*, JPL D-8451 (internal document), Jet Propulsion Laboratory, Pasadena, California, May 15, 1991.

JPL experimental data previously reported in the form of insertion loss and surface resistivity values. For this article, these experimental data were converted to electrical conductivity values. Data will also be presented on recently measured conductivity values of four test samples of the DSS-13 BWG steel shroud material.

II. Previous Test Data

Table 1 is a summary of all known previous experimental data on the electrical conductivity values of various types of steel at various microwave frequencies. In this table one can see that the experimental values at microwave frequencies range from a minimum of 0.0036×10^7 mhos/m to a maximum of 0.148×10^7 mhos/m, as compared with the dc value of 0.5×10^7 mhos/m. For purposes of comparison, the conductivity value of oxygen-free high-conductivity copper at microwave frequencies is 5.66×10^7 mhos/m, as compared with 5.8×10^7 mhos/m at dc. The term "effective conductivity" in Table 1 is the actual conductivity value divided by the relative permeability value. The relationships of permeability to surface resistivity and skin depth will be shown below.

From basic electromagnetic theory [8], it can be shown that the surface resistivity of a conductor in ohms/square is given as

$$R_s = 20\pi \sqrt{\frac{f_{\text{GHz}} \mu_r}{\sigma}} \quad (1)$$

where f_{GHz} is the frequency in GHz, μ_r is the relative permeability, and σ is the electrical conductivity in mhos/m.

By defining effective conductivity as

$$\sigma_{\text{eff}} = \frac{\sigma}{\mu_r} \quad (2)$$

Eq. (1) then becomes

$$R_s = 20\pi \sqrt{\frac{f_{\text{GHz}}}{\sigma_{\text{eff}}}} \quad (3)$$

If R_s is known at the measurement frequency but the effective conductivity is the parameter to be determined, one can use the following expression that was derived from algebraic manipulation of Eq. (3):

$$\sigma_{\text{eff}} = \left(\frac{20\pi}{R_s} \right)^2 f_{\text{GHz}} \quad (4)$$

where σ_{eff} is expressed in units of mhos/m.

The expression for skin depth in microinches can be derived as

$$\delta_{\mu i} = \frac{9972.58 R_s}{f_{\text{GHz}} \mu_r} \quad (5)$$

For the steel sample in Table 1, an effective conductivity of 0.0036×10^7 mhos/m was measured at 3.015 GHz. Since conductivity is approximately constant at microwave frequencies (except for surface-roughness and skin-depth effects), a calculation can be made by using this conductivity value for other microwave frequencies. Then for $f_{\text{GHz}} = 8.45$, from Eq. (3), $R_s = 0.9626$ ohms/square. If it is assumed that $\mu_r = 150$, then from Eq. (5), the skin depth is $7.57 \mu\text{in}$. However, if $\mu_r = 1000$, the skin depth becomes $1.14 \mu\text{in}$. For comparison purposes, it is of interest to make a similar calculation of a nonmagnetic material, such as aluminum, for $\mu_r = 1.0$ and a typical conductivity σ of 2.2×10^7 mhos/m. Then, at 8.45 GHz, $R_s = 0.0389$ ohms/square, and the skin depth is about $46 \mu\text{in}$.

It is known that at microwave frequencies, conductivity losses will increase with the increase of the ratio of surface roughness to skin depth [9]. Therefore, based on the previous discussion, one would expect the losses of a magnetic steel material with high permeability and a poor surface finish to have relatively high losses at microwave frequencies. As Vane points out, "Metals which have relatively high permeability are particularly lossy because the skin depth decreases with increasing permeability, thereby heightening the importance of surface irregularities." [1]

For interest, Table 2 is a tabulation of chemical composition and relative permeability values of stainless steel types 303, 304, 321, and 347. These are the types of stainless steel for which measured effective conductivity values are given in Table 1. It can be seen from Table 2 that the relative permeability values of these types of stainless steel are low (close to unity) and vary from 1.0–1.70. The relative permeability values of most types of stainless steel are functions of percent cold work reduction, but are generally less than 1.5. An exception is type 321 stainless steel whose relative permeability value has been measured to be as high as 9.4 for 71-percent cold work reduction and a magnetizing force of 200 Oe. Based upon these data, it can be stated that for most types of stainless steel materials,

the permeability will have little effect on the dissipative losses.

In contrast to the low magnetic properties of stainless steels, the relative permeability values of some types of carbon steel can be as high as 3000.² The relative permeability values of most structural or high-carbon-content steel are quite variable. If permeability of a particular type of steel must be known accurately, then the permeability values of samples of the particular material to be used need to be measured individually.

Since it was of interest to know the relative permeability of the DSS-13 shroud material, a sample of the material was given to the JPL Magnetism Test Laboratory for measurements of permeability. However, after several attempts using different methods, it was concluded that JPL does not currently have the equipment needed to measure permeability values of high-permeability steels.

III. Measurement Technique

A. Description of Technique

The TE₀₁₁ mode resonant-cavity technique employed to obtain the test results for this article was basically the same as that described previously [6,11], except that an HP 8510C Automatic Network Analyzer (ANA) was employed. Insertion losses at precisely known frequencies can be measured more accurately and more rapidly with the HP 8510C than with the nonautomated methods employed in the past. With the ANA, upon command, the computer finds the frequency of minimum insertion loss and the frequencies of the 3-dB points to 0.1-KHz resolution, as compared with 1-KHz resolution previously obtained with other setups. Insertion loss can be measured to 0.02-dB resolution. Visual displays of loss versus frequency are available on-screen, and an option is available for real-time hard copies of the frequency-response plots of the type shown in Fig. 1. The data can be stored for post-processing, later verifications and plotting, if desired. The application of the ANA for cavity measurements is considered to be a technical innovation that advances the state-of-the-art of reflector surface material evaluations.

B. Description of Shroud Test Samples

The DSS-13 shroud material is made from type ASTM 36 structural steel. As tabulated in [12], the chemical composition of this material, other than iron (Fe), is 0.26 per-

cent maximum carbon, 0.04 percent maximum phosphorus, and 0.05 percent maximum sulfur. A 24- by 24-in. section of steel shroud material was obtained for purposes of cutting it into smaller test samples for cavity measurements. This 24- by 24-in. section was a discarded piece that had previously been cut out of a bypass-mode rectangular shroud to permit clearance of a 29-dB horn installed at the bypass focal point. This cutout panel section was made from the same material as the BWG center-pass shroud and is 0.2 in. thick. After several 4- by 4-in. square test samples were cut from the 24- by 24-in. piece, slight deviations from flatness were found to exist on the test samples, and therefore the two possible test surfaces were labeled convex and concave. When a test sample was placed on the cavity opening (diameter = 2.24 in.) the maximum deviation from a flat plane was less than 0.004 in. For the particular measurement technique employed, it is not critical that the test samples be perfectly flat. Small deviations from perfect flatness affect primarily the resonant frequency. However, care was taken to identify which side of the test samples were actually measured.

IV. Test Results

The test results are shown in Table 3 and are tabulated according to types of materials. The results of most interest to the BWG project are those for (1) the BWG shroud samples' painted and unpainted sides, (2) 6061 T6 aluminum with and without the zinc-chromate primer, (3) 6061 T6 aluminum with and without the primer and Triangle no. 6 paint, and (4) galvanized steel samples.

An average effective conductivity value of 0.01 mhos/m for the unpainted BWG shroud samples agrees favorably with other steel conductivity values shown in Table 1. Of particular interest is the effect of paint on the steel surfaces of the BWG shroud. Triangle Co. thermal diffusive paint no. 6 made the effective conductivity significantly worse. The effective conductivity changed from the unpainted test-sample value of 0.01×10^7 mhos/m to an average painted surface value of 0.0036×10^7 mhos/m. From this result, one might conclude that if the shroud's steel surfaces were covered instead with high-conductivity paint, the resulting conductivity would be significantly improved. However, when one of the shroud samples had the paint removed and then was repainted with silver paint, as may be seen in Table 3, the effective conductivity value of the silver-painted sample was only a factor of two times better than that of the unpainted sample.

Primer and paint on aluminum 6061 T6 surfaces made the conductivity values only slightly worse than those of

² Calculated from carbon steel curve plotted among other dc magnetization curves for various magnetic materials published by the General Engineering Laboratory of the General Electric Company.

the unpainted aluminum surfaces. Data on the loss of primer and paint are difficult to obtain. The paint loss data, obtained from cavity measurements and presented in this article, are probably much more accurate than the data obtained from radiometric measurements [13]. Another interesting result to note in Table 3 is the measured conductivities of the galvanized steel samples. The two galvanized steel samples tested had conductivity values close to those measured for a brass test sample. This surprising result indicates that galvanized steel should be seriously considered as a candidate shroud material.

Also shown in Table 3 are the corresponding noise temperatures that would be generated if a plane wave were normally incident on an infinitely large flat sheet of the sample material. The noise temperatures were calculated from the approximate formula

$$NT \cong \frac{4R_s T_p}{\eta_o} \quad (6)$$

where

η_o = free space impedance (120π ohms)

T_p = physical temperature of the sheet (290 K)

It should be pointed out that these noise temperatures are not the same as the noise temperatures that would be generated due to conductivity losses when the material is used as the walls for a circular waveguide, such as a shroud.

V. Conclusions

Test data on the effective electrical conductivity of the DSS-13 shroud material have been presented. The painted shroud samples had an average effective conductivity of 0.0036×10^7 mhos/m. This conductivity value is about 140 times worse than the dc value of 0.5×10^7 mhos/m. The results of this study should prove useful in current investigations to determine the noise temperatures due to shroud dissipative losses. Other test data on various other types of steels and reflector surface materials have also been presented and should prove to be valuable for future reference purposes.

Acknowledgment

The authors thank Pablo Narvaez, of the JPL Reliability Engineering Section, for assistance in finding published information on the permeabilities of various types of steel.

References

- [1] A. B. Vane, "Measurement of Effective Conductivity of Metallic Surfaces at 3000 Megacycles and Correlation with Surface Conditions and DC Conductivity," Stanford Microwave Laboratory Report No. 4, Palo Alto, California, pp. III-9, 1949.
- [2] H. P. Westman, *ITT Reference Data for Radio Engineers (Fifth Edition)*, New York: Howard W. Sams, Inc., pp. 4-32, 1969.
- [3] T. Y. Otoshi, C. T. Stelzried, B. C. Yates, and R. W. Beatty, "Comparisons of Waveguide Losses Calibrated by the DC Potentiometer, AC Ratio Transformer, and Reflectometer Techniques," *IEEE Trans. on Microwave Theory and Techniques*, vol. MTT-18, no. 7, pp. 406-409, July 1970.

- [4] H. F. Reilly, J. J. Bautista, and D. A. Bathker, "Microwave Surface Resistivity of Several Materials at Ambient Temperature," *TDA Progress Report 42-80*, vol. October–December 1984, Jet Propulsion Laboratory, Pasadena, California, pp. 8–11, February 15, 1984.
- [5] T. Otoshi, "H-band Cryogenic Load Development Plating Evaluation," *JPL Space Program Summary 37-28*, vol. 4, Jet Propulsion Laboratory, Pasadena, California, p. 153, August 31, 1964.
- [6] R. C. Clauss and P. D. Potter, *Improved RF Calibration Techniques—A Practical Technique for Accurate Determination of Microwave Surface Resistivity*, Technical Report 32-1526, vol. 12, Table 1, Jet Propulsion Laboratory, Pasadena, California, p. 63, December 15, 1972.
- [7] A. C. Beck, "Conductivity Measurements at Microwave Frequencies," *Proceedings of the IRE*, vol. 38, no. 10, pp. 1181–1189, October 1950.
- [8] S. Ramo and J. R. Whinnery, *Fields and Waves in Modern Radio*, New York: John Wiley & Sons, 1950.
- [9] R. D. Lending, "New Criteria for Microwave Component Surfaces," *National Electronics Conference Proceedings*, vol. 11, pp. 391–401, 1955.
- [10] M. R. Gross, "Magnetic Permeability of So-Called Non-Magnetic Metallic Materials," *Journal American Science Naval Engineers*, vol. 66, no. 1, pp. 214–245, February 1954.
- [11] E. H. Thom and T. Otoshi, "Surface Resistivity Measurements of Candidate Subreflector Surfaces," *TDA Progress Report 42-65*, vol. July and August 1981, Jet Propulsion Laboratory, Pasadena, California, pp. 142–150, October 15, 1981.
- [12] American Society for Testing and Materials, *Annual Book of ASTM Standards*, "Standard Specification for Structural Steel," Designation A36/A 36M-90, vol. 01.05, Philadelphia, Pennsylvania, November 1990.
- [13] T. Y. Otoshi and M. M. Franco, "Radiometric Tests on Wet and Dry Antenna Reflector Surface Panels," *TDA Progress Report 42-100*, vol. October–December 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 111–130, February 15, 1990.

Table 1. Previously known or measured conductivity values of various types steel at dc and microwave frequencies.

Test sample	Surface roughness	Measurement method [Ref.]	Test freq., GHz	Effective conductivity, mhos/m	Square root of (dc cond./eff. cond.)
Steel, cold rolled	Machined	[1, Tables]	dc	0.50×10^7	
Steel (0.4–0.5% C, balance Fe)	Unknown	[2, Table 12, pp. 4–22]	dc	0.769×10^7 – 0.454×10^7	
Steel, cold rolled	Machined	Cavity [1]	3.015	0.0036×10^7	11.79
Stainless steel 430-2B	Rolled	Cavity [1]	3.015	0.0039×10^7	11.32
Type 303 stainless steel, test section 214	Electro-discharge machine (EDM) polished 8 μ in.	JPL insertion loss (IL) and National Bureau of Standards (NBS) Reflectometer [3]	2.115 2.295 2.388 2.3985	0.019×10^7	5.13
Type 303 stainless steel test section 215	EDM polished 20 μ in.	JPL insertion loss and NBS Reflect. [3]	2.295 2.388 2.3985	0.0136×10^7	6.06
Stainless steel, type 304	Polished <20 μ in.	Cavity JPL [4]	8.400	0.1233×10^7	2.01
Stainless steel type 347	Polished <20 μ in.	Cavity JPL [4]	8.400	0.0698×10^7	2.68
Mild steel	Polished <20 μ in.	Cavity JPL [4]	8.400	0.0301×10^7	4.08
Stainless steel type 321	Polished <20 μ in.	Cavity JPL [4]	8.400	0.0275×10^7	4.26
Stainless steel (type unknown)	Polished <20 μ in.	Cavity JPL [4]	8.400	0.0075×10^7	8.16
Type 304 WR112 stainless steel section part no. UP-1	Seamless cold drawn tubing, 70–100 μ in.	IL JPL [5]	8.448	0.0695×10^7	2.68
Type 304 WR112 stainless steel section UP-2	Same as UP-1 except inside electropolished, 20–35 μ in.	IL JPL [5]	8.448	0.1187×10^7	2.05
Type 304 WR112 stainless steel section UP-3	Same as UP-1 except inside finished by a perfect-peening process 35–45 μ in.	IL JPL [5]	8.448	0.1066×10^7	2.17
Stainless steel (type unknown)	Unknown	Cavity JPL [6]	8.415	0.00757×10^7	8.13
Iron wire	Unknown	Cavity Bell Laboratories [7]	9.000	0.148×10^7	1.84

**Table 2. Chemical compositions and relative permeabilities of various types of stainless steels at 24 deg C
(selected data from Tables IX and X of [10]).**

AISI ^a type stainless steel	Chemical composition, percent										Relative permeability for magnetizing forces of	
	C	S	P	Mn	Cr	Ni	Mo	Cu	Al	Other	H = 0.5 Oe	H = 100 Oe
303 annealed, machined	0.05	0.259	0.031	0.67	0.49	18.42	8.55	—	—	—	1.10	1.03
304 bar	0.073	0.030	0.037	1.03	0.47	18.87	8.47	0.30	0.24	—	<1.02	1.00
321 for plate thicknesses:												
5/8–1.0 in.	0.05	0.012	0.037	1.55	17.96	9.67	0.24	0.31	—	Ti 0.41	—	1.46–1.70
5/8 in.	0.05	0.006	0.032	1.22	17.35	9.83	0.22	0.18	—	Ti 0.39	—	1.06
5/8 in.	0.06	0.010	0.040	1.40	17.60	9.85	0.17	0.32	—	Ti 0.41	—	1.01–1.02
347 wrought pipe	0.08	0.010	0.020	1.69	18.13	12.34	—	—	—	Cb 0.80	<1.02	1.00

^a American Iron and Steel Institute.

Table 3. Test results of X-band cavity measurements.^a

Sample run ID	Run no.	Description	Side tested	Surface roughness, $\mu\text{in.}$	f_o , GHz	R_s , ohms/square	Effective cond., mhos/m	Flat sheet noise temp, K ^b
Silver	1-6	Cavity ref. plate			8.421883	0.0270 ^c	4.557×10^7	0.083
UP1CLR		BWG antenna shroud ASTM A36 steel sample no. 1, paint removed	Convex	230	8.420669	0.5993	0.0093×10^7	1.844
UP2CLR	1	BWG antenna shroud ASTM A36 steel sample no. 2, paint removed	Convex	>250	8.410031	0.5479	0.0111×10^7	1.686
UP2CLR	2	Same as above repeat test	Convex	>250	8.421861	0.5806	0.0099×10^7	1.787
UP3CLR		BWG antenna shroud ASTM A36 steel sample no. 3, paint removed	Concave	200	8.404988	0.5723	0.0101×10^7	1.761
UP4CLR		BWG antenna shroud ASTM A36 steel sample no. 4, paint removed	Concave	210	8.405010	0.5926	0.0095×10^7	1.823
UP1PNT		BWG antenna shroud ASTM A36 steel sample no. 1, plus Triangle 6 paint	Concave	134	8.406965	1.1184	0.0027×10^7	3.441
UP2PNT		BWG antenna shroud ASTM A36 steel sample no. 2, plus Triangle 6 paint	Concave	93	8.402868	0.9917	0.0034×10^7	3.051
UP4PNT		BWG antenna shroud ASTM A36 steel sample no. 4, plus Triangle 6 paint	Convex	156	8.403963	0.8353	0.0048×10^7	2.570
UP4SVR	1	BWG antenna shroud ASTM A36 steel sample no. 4, Triangle 6 paint removed and steel surface is now silver painted	Concave		8.407994	0.4913	0.0138×10^7	1.512
UP4SVR	2	Same as (UP4SVR, Run 1) above except silver paint allowed to dry longer	Concave		8.410720	0.3891	0.0219×10^7	1.197

^a Measurements were made on August 9, 1991.^b This is the noise temperature when the material is used as a flat plate short for normal incidence [see Eq. (6)]. This value is not the same as the noise temperature generated when this material is used as a BWG shroud.^c Average of six runs.

Table 3 (cont'd)

Sample run ID	Run no.	Description	Side tested	Surface roughness, $\mu\text{in.}$	f_o , GHz	R_s , ohms/square	Effective cond., mhos/m	Flat sheet noise temp, K ^b
COLDROLL		0.047-in.-thick cold rolled steel sheet	Side 1 Side 2	44 67	8.423088	0.3708	0.0242×10^7	1.141
COPPER	1	0.064-in. thick copper sheet	Side 1 Side 2	30 26	8.421679	0.0281	4.202×10^7	0.086
COPPER	2	Same as above except repeat test	Side 1 Side 2	30 26	8.420993	0.0275	4.388×10^7	0.085
COPPER	3	Same as above except repeat test	Side 1 Side 2	30 26	8.421060	0.0293	3.871×10^7	0.090
TYO1CLR	1	0.070-in. thick 6061-T6 Al sheet from DSS 13, bare metal side	Side 1	23	8.413483	0.0353	2.670×10^7	0.109
TYO1CLR	2	Same as TYO1CLR above except repeat test	Side 1	23	8.426800	0.0386	2.234×10^7	0.119
TYO1PMR	1	0.070-in. thick 6061-T6 Al sheet plus primer (primer is zinc-chromate side, ~ 0.6 mil)	Side 2 (other side of TYO1CLR)	64	8.409008	0.0378	2.321×10^7	0.116
TYO1PMR	2	Same as TYO1PMR above except repeat test	Side 2 (other side of TYO1CLR)	64	8.423660	0.0409	1.992×10^7	0.126
TYO2CLR	1	0.070-in. thick 6061-T6 Al sheet from GTS, bare metal side	Side 1	32	8.416702	0.0371	2.412×10^7	0.114
TYO2CLR	2	Same as above except repeat test	Side 1	32	8.426176	0.0381	2.289×10^7	0.117
TYO2PNT		0.070-in. thick 6061-T6 Al sheet primer plus Triangle no. 6 paint (measured thickness of primer plus paint ~ 1.1 mil)	Side 2 (other side of TYO2CLR)	89	8.420514	0.0439	1.722×10^7	0.135
TYO4CLR		0.063-in. thick 6061-T6 Al sheet bare metal side	Side 1	16	8.427346	0.0388	2.210×10^7	0.119
TYO4PNT		0.063-in. thick 6061-T6 Al sheet plus primer plus old Triangle no. 6 paint (measured thickness of primer plus paint ~ 1 mil)	Side 2 (other side of TYO4CLR)	26	8.425368	0.0399	2.086×10^7	0.123

Table 3 (cont'd)

Sample run ID	Run no.	Description	Side tested	Surface roughness, $\mu\text{in.}$	f_o , GHz	R_s , ohms/square	Effective cond., mhos/m	Flat sheet noise temp, K ^b
ALUM		0.064-in. thick 6061 Al sheet	Side 1	10	8.424538	0.0390	2.182×10^7	0.120
			Side 2	10				
BRASS		0.063-in. thick brass sheet	Side 1	14	8.424703	0.0492	1.372×10^7	0.151
			Side 2	21				
DB3S		0.024-in. thick galvanized steel sample	Side 1	22	8.426146	0.0526	1.201×10^7	0.162
			Side 2	23				
TYO3		0.048-in. thick, galvanized steel sample from DSS 13	Side 1	25	8.425480	0.0476	1.466×10^7	0.146
			Side 2	23				

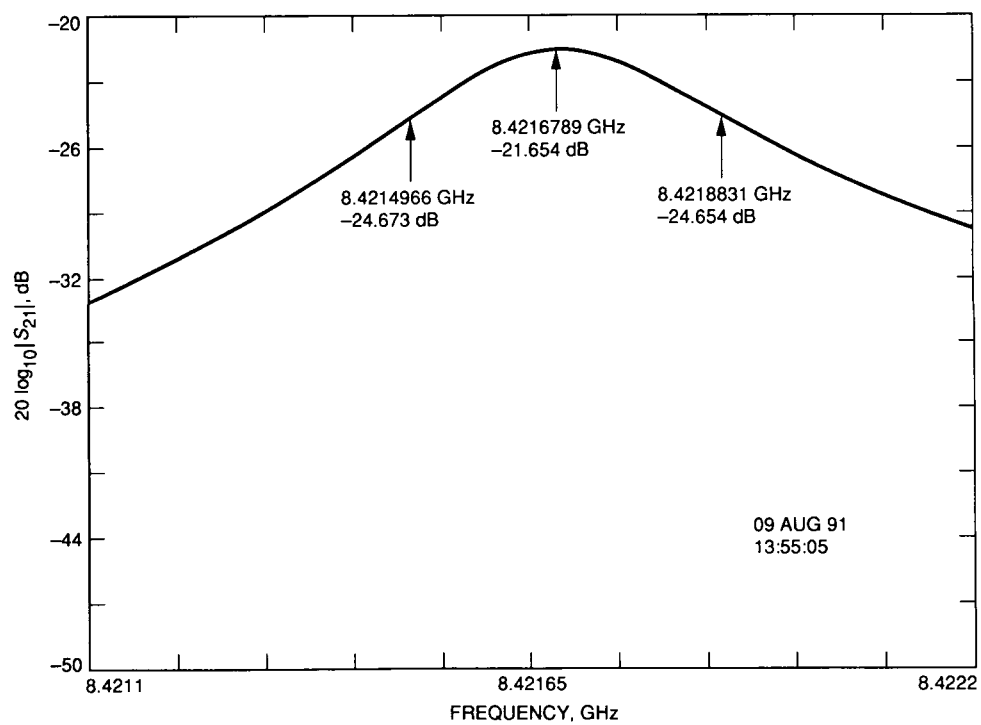


Fig. 1. Example of an automatic network analyzer plot of cavity-method insertion measurements on a copper test sample.

53-61
N92-24314-
P-21

Software-Engineering Process Simulation Model (SEPS)

C. Y. Lin

Information Systems Division

T. Abdel-Hamid

Naval Postgraduate School
Monterey, California

J. S. Sherif

Software Product Assurance Section
and
California State University at Fullerton

JJ 574450

NS 365219

JJ 574450

-CC 360121

This article describes the Software-Engineering Process Simulation (SEPS) model developed at JPL. SEPS is a dynamic simulation model of the software project-development process. It uses the feedback principles of system dynamics to simulate the dynamic interactions among various software life-cycle development activities and management decision-making processes. The model is designed to be a planning tool to examine trade-offs of cost, schedule, and functionality, and to test the implications of different managerial policies on a project's outcome. Furthermore, SEPS will enable software managers to gain a better understanding of the dynamics of software project development and perform postmortem assessments.

I. Introduction

The development and maintenance of software systems is a growing business; it has been estimated that U.S. expenditures for software development and maintenance were \$70 billion [14]. This figure is projected to grow to more than \$255 billion in 1995 [11], which accounts for 90 percent of total system expenditures [19]. However, the growth of the software industry has its downside. The record indicates that the development of software systems has been plagued by cost overruns, late deliveries, and dis-

satisfied customers [20,36,45,47,52]. These pervasive problems continue despite the significant software-engineering advances that have been made since the 1970s. There is a growing realization that most of these advanced technologies focus too much on solving the technical problems of software production and not enough on the managerial aspect of software project development [29,34,41].

In recent years, the management component has finally gained recognition as an area that is at the core

of both the problems, as well as the solutions for software crises [12,21,53]. For example, the defense-science board task-force investigation, led by Professor Frederick Brooks, concluded that the problems with military software development were not technical but managerial. The report urges not to apply technical fixes to what are management problems [15]. Furthermore, Dr. Richard F. Merwin, a guest editor for a special issue of *IEEE Transactions on Software Engineering* devoted to project management, pointed out that "Programming disciplines such as top-down design, use of standardized high-level programming languages, and program library support systems all contributed to production of reliable software on time, within budget What is still missing is the overall management fabric which allows the senior project manager to understand and lead major data processing development efforts." [34] And lastly, Abdel-Hamid notes, "Recently, it has become more and more evident that in software, product innovation is no longer the primary bottleneck to progress. The bottleneck is project management innovation." [3]

Thus, there is a growing belief that to minimize failures, good software-engineering practices are essential, but at the same time, there is a strong need to improve and advance existing software-management technology. Consequently, this has motivated the authors to develop a new software management technology—the Software Engineering and Management Process Simulation (SEPS) model, a software project-management tool.

The remainder of this section discusses software-management tools, the approaches used in existing tools, and a comparison of these tools with those of the SEPS model.

A. State of the Art

A number of different techniques and tools have been developed to aid software-development planning. There are cost-estimation and Critical Path Method/Program Evaluation and Review Techniques (CPM/PERT) models that support resource-management functions [9,13,42,49] and models for process-management functions [27,37,39, 54]. Each of these models has strengths and limitations. The strength of process models and CPM systems is their ability to model in great detail the activities within the software-development process. Their weaknesses include the inability to account for managerial decision making (such as a preference to hire versus reschedule) and the lack of feedback, among activities, that underlie software-development dynamics.

Cost-estimation models produce initial estimates that are essential for project start-up planning. These cost models, nevertheless, are limited. Most are static models (cost parameters are time-independent) designed to provide point estimates. They fail to capture management decision-making dynamics and their impact on project behavior. Furthermore, they are not well suited for real-time estimation adjustments once a project starts.

In summary, the existing management tools tend to focus on the software artifact and its transformation processes but fail to model the managerial and dynamic aspects that are at the core of software project-development.

B. SEPS Overview

SEPS is being developed to address the above weaknesses, specifically, it integrates technological aspects of software production with the managerial aspects. SEPS was designed so that it could be used to conduct trade-offs (on an ongoing basis) with regard to project cost, schedule, and functionality, and also would allow management to evaluate the implications of different managerial policies on a project's outcome. In addition, the authors sought to develop a model that would provide insight into the dynamics of the software-development process, since without a fundamental understanding of that process, the likelihood of any significant gains in software management is questionable [10,24,32].

To achieve that objective, the authors took a unique approach to developing SEPS that embodies: (1) dynamic feedback modeling, (2) an integrated view of the software project-development process, and (3) the use of simulation. The approach is articulated in more detail in [2]. The significance of these properties is summarized in the following paragraphs.

1. Dynamic Feedback Modeling. At the heart of the SEPS modeling task is the principle of dynamic feedback. The advantage of using dynamic feedback was described by Ondash, Maloney, and Huerta [40]. "A unique feature of dynamic project models not offered by networking planning methodologies is the ability to calculate the ripple (secondary) effects on project cost and schedule due to changing requirements. These changes might include changing government regulations . . . [or] work force availability. Ripple effects occur in labor productivity, unanticipated schedule slack and float time These processes can only be modeled by using dynamic modeling with explicitly represented feedback mechanisms. In this respect, dynamic project models complement the static . . . mod-

els by providing the capability to readily perform sensitivity analyses of likely perturbations and their consequential ripple effects.”

2. An Integrated View. Not only does SEPS encapsulate and simulate the dynamic interactions among various software life-cycle engineering activities (e.g., design, rework, inspection), but also it illustrates the influential relationships among management functions (planning, controlling, monitoring) and engineering activities (see Fig. 1). Only the integration of both engineering and management aspects allows one to examine the effectiveness of an intended management decision on overall project performance [2,31].

3. Simulation. Simulation is a powerful technique used to handle the complexity of the model (hundreds of dynamic variables and causal linkages). The use of simulation enables managers to assess quickly and safely the implications of an intended policy before it is implemented. Also, users can easily conduct controlled postmortem experiments to develop new insights into a project.

In summary, employing a feedback-modeling technique allows one to capture the dynamic ripple effect that characterizes software project development, and the integrated approach enforces an explicit description and enhances the understanding of the interrelationships among engineering activities and managerial decisions. Finally, the use of simulation helps to reduce uncertainties and risks associated with a policy and provides support for ongoing project replanning.

Section II provides a summary of the process used to develop SEPS and its overall model structure. A detailed description of a submodel structure and some of its mathematical formulations are also discussed.

II. SEPS Model Development

The SEPS feedback structure was created by a rigorous process that included three critical development steps: field interviews, literature review, and peer/expert review. The first step was to conduct a series of interviews with various software managers and engineers. Their views and hands-on experiences of how software systems are produced were used to develop SEPS’s core feedback structure. The information gathered included management and engineering practices, strategies, activity interactions, and the influential relationships among managerial decisions and development activities. After this information was incorporated into the model structure, a literature review

was performed. The literature review provided the following benefits:

- (1) It verified the feedback structure by checking the structure obtained from field interviews with observations from the open literature.
- (2) It supplemented knowledge in areas that are closely related to software development (such as management control, psychology, and organization behavior), and therefore enabled the authors to enhance the overall model structure [2].

At the end of the literature-review step, a fairly comprehensive and integrated software project-development model was produced. This model was then subjected to iterative review and critique by experts (from JPL and other NASA centers) in the areas of modeling and simulation, software project management, and software engineering. The review process produced a model structure that closely mimics the software project-development process.

The following sections present an overview of the SEPS model feedback structure. Included in the discussion are examples of mathematical formulations used in the model.

A. SEPS Structure Overview

The SEPS model consists of the following four submodels: Production, Staff/Effort, Scheduling, and Budget. Each submodel represents one part of the software project-development process and is linked to the others by a management-decision network. Together, these components represent an integrated view of the dynamics of the software project-development process. Figure 2 depicts a high-level view of the relationships among the four submodels.

The Production submodel captures the various software production activities, their dependencies and interrelationships, and the functions that determine work progress. The Staff/Effort submodel simulates the functions that determine required work-force levels and mimics the flow of personnel resources. The Scheduling submodel encapsulates the functions that determine the time to complete a task and forecasts a completion time for each software life-cycle phase. The Budget submodel calculates expenditures and accumulated costs.

The decision-making characteristics in SEPS are derived from the dynamic feedback of information among the planning (e.g., for resources), monitoring (e.g., for product development), and controlling (e.g., the development process) functions. Also, since SEPS encompasses each

life-cycle phase, the structure shown in Fig. 2 exists within each phase.

Figure 3 illustrates a second level of abstraction for the interrelationships among the various components within each submodel. This figure shows how the submodels interrelate through a set of information-feedback links.

It is beyond the scope of this article to discuss in detail the entire SEPS model structure. Instead, the next section provides examples of the feedback structure of the variables included in the Production submodel and the relationships of these variables to other submodels. The relationships among variables are described by mathematical formulations.

B. SEPS Feedback Structure: An Example

The Production submodel starts with a number of tasks (input parameter) to be developed. The number of tasks uncompleted is normally depleted through the software-production rate. However, as a project continues to develop, its scope is frequently altered. The changes can be attributed to various sources, such as the discovery of new tasks when requirements are better defined, or descoping tasks due to schedule or budget overruns. This relationship is illustrated in the upper left-hand quarter of Fig. 3.

Let $W(t)$ denote the number of tasks to be developed at any time t (i.e., the backlog of work). Furthermore, let $P(t - dt, t)$, $R(t - dt, t)$ and $N(t - dt, t)$ denote the team production rate (during time interval $t - dt, t$), task descoping rate, and new task discovery rate, respectively. There exists a relationship such that

$$\begin{aligned} W(t) &= W(t - dt) \\ &+ dt[N(t - dt, t) - P(t - dt, t) - R(t - dt, t)] \end{aligned} \quad (1)$$

where dt is the time-increment interval. To simplify the discussion, the authors treat $R(t - dt, t)$ and $N(t - dt, t)$ as exogenous variables that mimic management decisions based on information generated from other parts of the system (as shown in Fig. 3). These parameters can be expressed as a pulse function or a time-tagged table function. The team production rate is defined as $P(t - dt, t)$, which is a function of staff size, $S(t)$; average productivity rate, $P_a(t)$; intercommunication-overhead factor, $C(t)$; learning factor, $L(t)$; and work-intensity factor, $F(t)$. The team production rate takes the form

$$P(t - dt, t) = [\Sigma S_i(t)] * P_a(t) * C(t) * L(t) * F(t) \quad (2)$$

where each $S_i(t)$ is the number of full-time equivalent staff of type i that is allocated for production at time t . In the model, staff is classified by their origins and experience levels. This results in four distinct groups of S_i : in-house senior staff, in-house junior staff, newly hired senior staff, and newly hired junior staff. $S_i(t)$ is modeled in the Staff submodel as

$$\begin{aligned} S_i(t) &= S_i(t - dt) \\ &+ dt[A_i(t - dt, t) - T_i(t - dt, t) - X_i(t - dt, t)] \end{aligned} \quad (3)$$

where $A_i(t - dt, t)$ is the staff-assignment rate, $T_i(t - dt, t)$ is the staff-release rate, and $X_i(t - dt, t)$ is staff-assimilation rate.

The weighted average productivity rate, $P_a(t)$ is defined as

$$P_a(t) = \frac{\Sigma(S_i(t) * P_i)}{\Sigma S_i(t)} \quad (4)$$

where each P_i is the nominal (unencumbered) individual staff-productivity rate for staff type i . Here the authors encapsulate within P_i (a constant for each staff type) those productivity determinants cited by Boehm [13]. The reason for using static nominal productivity is based on the observation that even though productivity determinants vary from organization to organization, and from project to project within an organization, they remain constant within a single project [2].

There are three productivity determinants that exhibit dynamic characteristics during project development and, therefore, their effects on productivity are modeled explicitly. The determinants are the intercommunication overhead factor, $C(t)$; the learning factor, $L(t)$; and the actual fraction of time an individual spends on a project, or work-intensity factor, $F(t)$.

The intercommunication overhead factor, $C(t)$, is cited by many software researchers and practitioners as a significant productivity determinant [2]. It is generally agreed that increases in staff level for a project could have a negative effect on team productivity [2,16,48,55]. The reason for this behavior is that adding staff to a project creates

additional communication overhead, such as verbal communication, documentation, and interfaces, all of which cause each team member's productivity to drop below his or her normal rate. In Tausworthe's work on intercommunication overhead [50], he points out that the team productivity rate is affected by the intercommunication overhead factor, $C(t)$, such that¹

$$C(t) = [1 - t(S)]$$

where $t(S)$ is the relative productivity loss due to intercommunication among staff and is a function of staff size S . In the SEPS model, $t(S)$ is derived from the data gathered from field interviews and the open literature [2], and it is defined as

$$t(S) = 1 - [1.03 \exp(-0.02S)]$$

and, therefore,

$$C(t) = 1.03 \exp(-0.02S) \quad (5)$$

The learning factor, $L(t)$, can again be modeled by a time-tagged table function in which the independent variable is the percentage of work completed. The parametric value of $L(t)$ shown in Fig. 5 has an S-shaped curve, as suggested by several authors [8,17,53].

Finally, the authors define $F(t)$ as the work intensity, or the actual fraction of time an individual spends working on a project. In his doctoral work at the Massachusetts Institute of Technology, Professor Abdel-Hamid conducted extensive research in this area. His review of the literature [4,5,6,25,30,33], combined with field interviews (5 organizations), led him to conclude that in the absence of schedule pressure, a person devotes, on average, 60 percent of his or her time working on project-related tasks. The remaining "slack time" is consumed by miscellaneous activities not directly related to the project, such as conferences, personal telephone calls, coffee breaks, and so on [2]. However, the effective work fraction fluctuates throughout a project. It would decrease even more if a project were ahead of schedule (negative schedule pressure), or increase, potentially up to 100 percent [13], if a project were behind schedule. Since it is known that $F(t)$ can fluctuate dynamically throughout a project's development, one can express it as a state variable

$$F(t) = g(t) + [F(0) - g(t)] \exp(-t/T) \quad (6)$$

or

$$F(t) = F(0) \left(1 + x(t) [1 - \exp(-t/T)] \right)$$

where $F(0)$ is the initial value of $F(t)$ (e.g., 0.6), and T is the constant time required to realize the need to adjust work intensity. The desired fraction of daily work hours at time t is denoted by $g(t)$ and is expressed as

$$g(t) = (1 + x(t)) * F(0)$$

where $x(t)$ is the fractional increase or decrease of effective work time, based on schedule pressure, from the initial fraction (i.e., 0.6).

Errors are made during the software development process; this characteristic is defined in SEPS as the error-generation rate. These human errors produce defects in products that remain undetected until product quality-assurance (QA) activities occur (e.g., formal inspections, design/code walk-throughs, and peer reviews). The QA effort will discover some defective products, but others will escape detection. These undetected faulty products will further affect the quality of subsequent products. This phenomenon is called the error-explosion factor. Let $E(t - dt, t)$ denote the error-generation rate, defined as

$$E(t - dt, t) = [P(t - dt, t) * N(t) * Y(t) * M(t)] \\ + [P(t - dt, t) * U(t) * Z] \quad (7)$$

The first part of the right-hand side of Eq. (7) indicates that the defects are generated proportionally to the team production rate, $P(t - dt, t)$. This is defined by the relationship between $N(t)$, the nominal number of defects committed per unit of work (e.g., source line of instruction), and the dynamic factors $Y(t)$ and $M(t)$, where $Y(t)$ is the schedule pressure factor and $M(t)$ is the work-force mix factor. $N(t)$ can be modeled through a table function in which the independent variable is the percentage of work completed for a given task, as shown in Fig. 6. The range of $N(t)$ parameter values is based on the empirical evidence gathered by Boehm [13]; Jones [26]; and Thayer, Lipow, and Nelson [51]; however, these values are project dependent and can be calibrated accordingly. Notice that $N(t)$ is a nominal value, which means that a

¹ The authors took the liberty of changing Tausworthe's notation so that it would be consistent with the notation used in this article.

certain number of errors will be created even though a project is right on track (e.g., no schedule slippage or unplanned workforce fluctuation). However, many software projects do experience some sort of schedule pressure (also known as schedule crunch) during their development life cycle. Schedule pressure has been found to cause an increase in the number of errors generated [1,4,5,35,43,44]. Hence, SEPS captures this phenomenon with variable $Y(t)$ as the error-generation modifier due to schedule pressure, and its parametric value is shown in Fig. 7.

Finally, $M(t)$, the workforce-mix factor, is the last modifier of the first part of the right-hand side of Eq. (7). When a project is facing a schedule crunch, management is often inclined to add staff to get the project back on track. Even though it would be desirable to add experienced in-house staff (e.g., a tiger team), these people may not be available. In such a case, management must find less-experienced in-house personnel or hire staff from outside the organization. In either approach, there would be an impact on both productivity and error generation (increase) [18,38]. SEPS introduces $M(t)$ to capture the dynamic changes in the workforce-mix ratio (e.g., experienced versus in-experienced staff) and its impact on the error-generation rate. The parametric value of $M(t)$ is shown in Fig. 8 (derived from data gathered by Albrecht [5], Artzer, and Neidrauer [7]).

Thus far, error generation has been addressed with the assumption that the previous life-cycle-phase output product used to create the current-phase product is flawless. Of course, this is often not the case. One must therefore consider the existence of undiscovered defects from previous life-cycle phases that create additional defects downstream. The second part of the right-hand side of Eq. (7) is designed to model this characteristic. The behavior is modeled by saying that during production, additional defects are created due to the density of previously undetected defects, $U(t)$, with an explosion factor, or latency, Z . $U(t)$ is defined as

$$U(t) = \frac{(\text{Undiscovered defects from previous phase})}{\int (P(t - dt, t) dt \text{ from previous phase})} \quad (8)$$

and the latency factor Z accounts for the fact that, for example, one design error will explode into five coding errors. Another way to express this behavior is that it will require more effort (often substantially more) to fix an error that was not discovered in a previous phase. In SEPS, the Z 's are constants (since SEPS is a life-cycle model, Z is allowed to have a distinct constant for each life-cycle

phase). The constant values of Z used in SEPS are based on the empirical study conducted by Kelly and Sherif [28]. In that study, the researchers found that the explosion factor for developing ground software using Ada and/or C is about 15, from requirements analysis to coding. For flight software, the explosion factor from requirements analysis to coding is about 25.

Finally, a look at the error-detection rate completes the discussion of the Production submodel. As noted earlier, defective products remain undiscovered until QA activities occur. The causal loop structure for the error-detection rate, $D(t)$, shown in Fig. 3, is illustrated in Fig. 4. Specifically, the error-detection rate is a function of the product-inspection rate, $I(t - dt, t)$; the error density, $\delta(t)$; and the inspection efficiency factor, $\lambda(t)$. This relationship is defined as

$$D(t - dt, t) = I(t - dt, t) * \delta(t) * \lambda(t) \quad (9)$$

and

$$I(t - dt, t) = \frac{-dW(t)}{dt} \quad (10)$$

where $W(t)$ denotes work units that have been developed and are ready for inspection, and

$$\delta(t) = \frac{PD(t)}{w(t)} \quad (11)$$

where PD indicates the level of potential detectable faulty products and is defined as

$$PD(t) = PD(t - dt) + dt[E(t - dt, t) - I(t - dt, t) * \delta(t)] \quad (12)$$

Empirical studies suggest that the effectiveness of detecting faulty products has a strong correlation to the level of QA effort allocated and the amount of products to be inspected in a given period of time (this correlation exists regardless of the QA technique used). The parametric values used in SEPS to describe $\lambda(t)$ correspond with these studies and are shown in Fig. 9. The behavior exhibited in Fig. 9 indicates that the efficiency indicators correspond to the dependent variable $J(t)$, the units of work that are being inspected per staff hour (or week), as defined by

$$J(t) = \frac{I(t - dt, t)}{S_q(t)} \quad (13)$$

where $S_q(t)$ is the actual level of staff allocated to QA, and is computed in the Effort/Staff submodel. Basically, $S_q(t)$ is derived from the planned fraction of effort allocated for QA, a management-decision input parameter, with schedule pressure and training factors impacting the planned QA effort fraction.

So far, the SEPS model has been reviewed, and its development and feedback structure has been explained. In Section III, the validity of SEPS is demonstrated by reviewing its behavior when subjected to sensitivity analysis, and comparing it with the actual characteristics of a real project.

III. Validation of the SEPS Model

A battery of validation tests designed to evaluate the SEPS model is being carried out in two separate phases. Phase-1 validation tests were conducted in FY'91, and phase-2 validation tests will be performed during FY'92. Tests conducted in phase 1 included sensitivity analysis and historical project comparison. Sensitivity tests are used to check how sensitive the model is to perturbations over a range of input parameters, whereas historical project-comparison tests validate how precisely the model replicates historical projects. The results of phase-1 tests are presented in the following sections.

A. Sensitivity Analysis Tests

To conduct this test, the authors created a hypothetical software project where the project size is 128,000 source lines of code (sloc). The initial effort estimate was 1621 work-weeks with 95 weeks of project duration and an average staff size of approximately 18 people.

Sensitivity analysis tests over nine SEPS input parameters were performed. The input parameters are:

- (1) Staff experience level.
- (2) Project size.
- (3) Initial estimated schedule (compression/relaxation).
- (4) Quality-assurance effort.
- (5) Initial staff size.
- (6) Staff-size limit.
- (7) Initial estimated effort.
- (8) Error-generation rate.

- (9) Annual turnover rate.

The graphical results of the sensitivity tests are shown in Figs. 10–18. The data were normalized for ease of demonstration of the results. A brief description of each test follows

1. Staff Experience Level. A project normally consists of people with mixed experience levels; in this experiment the software engineers are grouped into two categories: experienced and inexperienced staff. To demonstrate the effect of average team skill on the project (in terms of experience), the percentage of experienced staff level was varied from 30 to 100 percent, with 80 percent as the base case. The results indicate that *as the fraction of experienced staff increases, the total effort and completion time required to complete the project decreases; however, the impact on project completion time is less significant than it is on effort. This is due to the management pressure to finish the project on schedule.* (See Fig. 10.)

2. Project Size. It is well publicized that there is a strong correlation between the size of a software project and the effort and schedule needed to complete a project. In this experiment the authors verified whether SEPS can produce such behavior by varying the project size from 76.8 ksloc to 192 with 128 ksloc, as the base case. The results show that *as the project size increases, both the total effort and completion time increase.* (See Fig. 11.)

3. Initial Estimated Schedule (Compression/Relaxation). Studies have demonstrated that the schedule estimated at the beginning of a project has a profound impact on the project's outcome. To investigate the effect the authors gradually compressed the initial estimated schedule from the base case (95 weeks) down to 60 percent (58 weeks) and also relaxed it to 150 percent (142 weeks). The results show that *as the initial estimated schedule is compressed, the total effort required increases and, on the other hand, as the estimated schedule is extended, the total effort required decreases slightly.* (See Fig. 12.)

4. Quality Assurance (QA) Effort. In this experiment the authors studied the effect of the percentage of total effort allocated to the software QA activities on project outcomes. The percentage of the QA effort was varied from 5 percent to 45 percent, with 15 percent as the base case, and the results show that *both total effort and completion time increase if the planned quality assurance effort is either too high or too low.* (See Fig. 13.)

5. Initial Staff Size. The initial staff size is defined in this experiment as the number of people assigned at the

beginning of a project. By using the Constructive Cost Model (COCOMO) [11], an average staff size of 8 and peak staff size of 10 were obtained for the requirements-analysis phase. The authors started the experiment with 2 people and gradually increased the number to 10, with 8 as the base case. The results indicate that *as the initial requirements staff size is increased (up to the planned peak value for the requirements phase), both total effort and completion time decrease*. (Keep in mind that the authors assume that the staff will continue to participate throughout project life-cycle development; see Fig. 14.)

6. Staff Size Limit. The workforce limit is defined as the maximum allowable peak staff size during project development. In this experiment, the authors first ran the model without a workforce cap. As a result, the staff level peaked at 25. Next, the authors varied the workforce limit from 10 to 25, with 20 as the base case. The experiment results demonstrate that *as the staff size limit is decreased (due to a hiring freeze, for example), the project completion time is increased and the total effort is decreased; however, the drop in total effort is not significant*. (See Fig. 15.)

7. Initial Estimated Effort. The initial estimated effort is defined as the total effort, estimated at the beginning of the project, required for developing a project. In this study the authors examined the effect of the initial estimated effort variation on the project's final total effort and completion time. The results show that *the project's final total effort increases when the initial estimated effort is either severely underestimated or overestimated. The completion time tends to decrease as the initial estimated effort increases*. (See Fig. 16.)

8. Error-Generation Rate. As explained in Section II, during the product-development phase, it is almost certain that errors will be generated. In this experiment certain error rates are assumed and their effects on the project outcome is examined. When the error rates were varied from 2 to 40, with 10 (errors per ksloc) as the base case, the results show that *the higher the error generation rate is, the higher the total effort and the longer the completion time. It is assumed that the planned allocated QA effort is a fixed function*. (See Fig. 17.)

9. Annual Turnover Rate. It is a management nightmare to be confronted with a large staff-turnover rate. In this study the impact of varying staff annual turnover on project total effort and completion time is examined. The results show that *a moderate increase in annual turnover rate causes an increase in completion time and no visible*

variation in total effort. A higher turnover rate causes both the total effort and completion time to increase. (See Fig. 18.)

B. Evaluation of Sensitivity-Analysis Tests

Sensitivity analysis involves determining how much the simulation output will vary with a small change in an input parameter. It is a tool used to characterize the unique features of the original model. In general, sensitivity analysis has four main uses: (1) assessing and interpreting the reasonableness of simulation, (2) experimental exploration of the model, (3) better allocation of resources for further data collection, and (4) promoting model simplification [46].

To evaluate the sensitivity-analysis tests, the authors relied on the experience of project managers and researchers at JPL and Goddard Space Flight Center (GSFC). Twenty-three staff members with an average of more than fifteen years experience in software management and/or costing evaluated the sensitivity-analysis tests. Their assessments of the comparative accuracy of the model output to real life are summarized in Table 1.

The stratified survey of 23 experienced managers and researchers at JPL and GSFC rated the behavior of the simulation model as reasonable 88 percent of the time, with a standard deviation of 0.02. A 95-percent confidence interval for the overall evaluation extends from 84 to 92 percent.

C. Historical Project Comparison

The sensitivity test is very effective in revealing abnormalities in the model-generated behavior, which in turn indicate possible problems with the model structure or mathematical equations. In essence, this test examines a model's correctness. However, the test does not confirm a model's accuracy in prediction. Another test was therefore conducted—a historical project case comparison—to validate the SEPS prediction capability.

The Cosmic Background Explorer (COBE) Attitude Ground Support System (AGSS) project at GSFC was chosen for the study as it maintained a detailed database and a well-documented report that describes the project-development history. The COBE/AGSS system was designed and implemented to support the COBE spacecraft mission, which began in July 1986 and was completed in August 1988. Its functional requirements include [23]:

- (1) Provide ground-attitude determination.

- (2) Monitor and verify attitude-control system (ACS) performance.
- (3) Provide attitude-sensor alignment and calibration.
- (4) Provide spacecraft attitude-control support.
- (5) Provide ACS prediction support.
- (6) Provide testing and simulation support.
- (7) Provide contact prediction support.

The project characteristics are summarized as follows (see [23] for a detailed description of the project):

- (1) Estimated size: 94.1 ksloc; actual size: 163.2 ksloc.
- (2) Estimated effort: 241 staff months; actual effort: 336 staff months.
- (3) Estimated duration: 85 staff weeks; actual duration: 110 staff weeks.
- (4) Language used: mostly Fortran, some Assembler.
- (5) Incremental development process used: three implementation builds and two releases.
- (6) Computer Science Corporation (CSC) was the contracting developer.
- (7) Project started in July 1986 and ended in August 1988.

The COBE/AGSS life-cycle phases modeled by SEPS include:

- (1) Requirements analysis.
- (2) Preliminary design.
- (3) Detailed design.
- (4) Implementation:
 - (a) Three developmental builds.
 - (b) Three build integration/tests.
- (5) System test.

The initial COBE/AGSS project estimates used as input parameters to SEPS include:

- (1) Initial project size (94.1 ksloc).
- (2) Initial project effort (241 staff months).
- (3) Planned life-cycle effort distribution [22].
- (4) Initial schedule estimates for the life cycle modeled.
- (5) Project-size growth estimate.
- (6) Nominal error-generation rate.

There are approximately twenty other parameters (e.g., hiring delay, nominal staff productivity, nominal staff attrition) that are calibrated to the COBE/AGSS development environment.

The validation scheme is presented in Fig. 19. The SEPS-predicted versus COBE-actual project life-cycle staffing curve over time is shown in Fig. 20. The root mean squared error (RMSE) as a measure of accuracy between the predicted values of SEPS and actual values of COBE is 1.77 persons. Analysis of variance shows that on the average there is no significant difference between the values predicted by SEPS and actual project life-cycle staffing values of COBE at the level of significance of $\alpha = 0.01$. Figure 21 illustrates a comparison of the accumulated effort predicted by SEPS versus actual accumulated effort as a function of time. The RMSE measure of accuracy between the predicted values of SEPS and actual values of COBE is 13.6 person weeks. Analysis of variance shows that on the average there is no significant difference between the values predicted by SEPS and COBE's actual project-accumulated effort at the level of significance of $\alpha = 0.01$. Finally, the schedule comparisons for each life-cycle phase are given in Fig. 22. The RMSE measure of accuracy between the values predicted by SEPS and the actual values of COBE is 2 weeks. Analysis of variance shows that on the average there is no significant difference between the values predicted by SEPS and COBE's actual project schedule end date at the level of significance of $\alpha = 0.01$.

IV. Conclusions

The increasing awareness of the need to improve the quality of managerial software motivates the software industry to come up with better management techniques and tools.

This article briefly reviewed some existing tools and then focused on the SEPS model. SEPS was developed to support software project planning and prevent software project-development failures. The specific objectives of SEPS are to assist software managers in preproject contingency analyses and support project replanning (of cost and schedule, for example) during the development life cycle. In addition, SEPS provides a learning environment through simulation where the implications of different policies on a project can be studied, and insight can be gained into the causes of project dynamics.

Although more testing needs to be conducted, the findings from the sensitivity test—with a confidence rating of

C-3

88 percent from the evaluators, and the results from the COBE historical project comparison at the level of significance $\alpha = 0.01$ —give the researchers and evaluators great confidence in the validity of the SEPS model.

SEPS has demonstrated its ability to replicate the software project dynamics observed in the software industry, and a specific project at GSFC, the next challenge for SEPS is to validate its applicability to the DSN.

References

- [1] T. K. Abdel-Hamid, "On the Utility of Historical Project Statistics for Cost and Schedule Estimation," *Journal of Systems and Software*, vol. 13, pp. 71-82, 1990.
- [2] T. K. Abdel-Hamid, *Software Project Dynamics: An Integrated Approach*, Englewood Cliffs, New Jersey: Prentice-Hall, 1991.
- [3] T. K. Abdel-Hamid, "Organizational Learning—The Key to Software Management Innovation," *American Programmer*, vol. 4, no. 6, pp. 20-27, June 1991.
- [4] D. S. Alberts, "The Economics of Software Quality Assurance," *Proceedings of the National Computer Conference*, Montvale, New Jersey, pp. 433-442, 1976.
- [5] A. J. Albrecht, "Measuring Application Development Productivity," *Proceedings of the Joint SHARE/GUIDE/IBM Application Development Symposium*, New York, pp. 83-92, October 1979.
- [6] R. N. Anthony, *Planning and Control Systems: A Framework for Analysis*, Cambridge, Massachusetts: Harvard University Press, 1979.
- [7] S. P. Artzer and R. A. Neidrauer, "Software Engineering Basics: A Primer for the Project Manager," Unpublished thesis, Naval Postgraduate School, Monterey, California, 1982.
- [8] J. M. Buxton, P. Naur, and B. Randell, eds., *Software Engineering: Concepts and Techniques*, New York: Litton Educational Publishing, Inc., 1976.
- [9] J. W. Bailey and V. R. Basili, "A Meta-Model for Software Development Resource Expenditures," *Proceedings of the 5th International Conference on Software Engineering*, IEEE/ACM/NBS, New York, pp. 107-116, March 1981.
- [10] V. R. Basili, *Tutorial on Models and Metrics for Software Management and Engineering*, New York: Computer Society Press, 1980.
- [11] B. W. Boehm, "Improving Software Productivity," *Computer*, vol. 20, pp. 43-50, September 1987.
- [12] B. W. Boehm, *Software Risk Management*, New York: IEEE Computer Society, 1989.
- [13] B. W. Boehm, *Software Engineering Economics*, Englewood Cliffs, New Jersey: Prentice-Hall, 1981.
- [14] B. W. Boehm and P. N. Papaccio, "Understanding and Controlling Software Costs," *IEEE Transactions on Software Engineering*, vol. 14, no. 10, pp. 1462-1477, October 1988.
- [15] F. P. Brooks, Jr., "IEEE Software," *Softnews*, vol. 13, p. 87, January 1988.
- [16] F. P. Brooks, Jr., *The Mythical Man-Month*, Reading, Massachusetts: Addison-Wesley Publishing Co., 1978.

- [17] J. Cooper, "Software Development Management Planning," *IEEE Transactions on Software Engineering*, vol. 10, pp. 22-26, 1984.
- [18] A. G. Endres, "An Analysis of Errors and Their Causes in System Programs," *IEEE Transactions on Software Engineering*, vol. 1, pp. 140-149, June 1975.
- [19] R. E. Fairley, *Software Engineering Concepts*, New York: McGraw-Hill Book Co., 1985.
- [20] W. L. Frank, *Critical Issues in Software: A Guide to Software Economics, Strategy, and Profitability*, New York: John Wiley and Sons, 1983.
- [21] P. F. Gehring, Jr. and V. W. Poovch, "Software Development Management," *Data Management*, vol. 8, pp. 14-38, February 1977.
- [22] *Manager's Handbook for Software Development*, Revision I, Goddard Space Flight Center, Baltimore, Maryland, November 1990.
- [23] *Software Development History for Cosmic Background Explorer (COBE) Attitude Ground Support System (AGSS)*, Goddard Space Flight Center, Baltimore, Maryland, December 1988.
- [24] W. S. Humphrey, *Managing the Software Process*, Reading, Massachusetts: Addison-Wesley Publishing Company, 1989.
- [25] R. L. Ibrahim, "Software Development Information System," *Journal of Systems Management*, vol. 11, pp. 34-39, December 1978.
- [26] T. C. Jones, "Measuring Programming Quality and Productivity," *IBM Systems Journal*, vol. 17, no. 1, pp. 39-63, 1978.
- [27] M. Keller and M. Wilkins, "On the Use of an Extended Relational Model to Handle Changing Incomplete Information," *IEEE Transactions on Software Engineering*, vol. 11, no. 12, pp. 620-633, 1985.
- [28] J. Kelly and Y. S. Sherif, "Where Is The Most Critical Point to Emphasize Software Inspections?," *IEEE Transactions on Software Engineering*, submitted in 1991.
- [29] K. Kolence, "Software Engineering Management and Methodology," *Software Engineering: Report on a Conference Sponsored by the NATO Science Committee*, edited by P. Naur and R. Randell, New York: IEEE Press, vol. 13, pp. 116-123, October 1968.
- [30] J. H. Lehman, "How Software Projects Are Really Managed," *Datamation*, vol. 25, pp. 35-41, 1979.
- [31] C. Y. Lin and R. Levary, "Computer-Aided Software Process Design," *IEEE Transactions on Software Engineering*, vol. 15, no. 9, pp. 1025-1037, September 1989.
- [32] L. Liu and E. Horowitz, "A Formal Model for Software Project Management," *IEEE Transactions on Software Engineering*, vol. 15, no. 10, pp. 1280-1293, October 1989.
- [33] C. L. McGowan and R. C. Henry, "Software Management" in *Research Directions in Software Technology*, edited by P. Wegner, Cambridge, Massachusetts: MIT Press, 1979.
- [34] R. E. Merwin, "Software Management: We Must Find a Way," *IEEE Transactions in Software Engineering*, no. 4, pp. 307-361, 1978.

- [35] H. D. Mills, *Software Productivity*, Toronto, Canada: Litter, Brown & Co., 1983.
- [36] S. N. Mohanty, "Software Cost Estimation: Present and Future," *Software—Practice and Experience*, vol. 11, pp. 103–121, 1981.
- [37] P. Montgomery, Jr., "A Model of the Software Development Process," *Journal of Systems and Software*, vol. 2, pp. 237–255, 1981.
- [38] C. J. Myers, *Software Reliability: Principles and Practices*, New York: John Wiley and Sons, 1976.
- [39] M. Newman and R. Sabherwal, "A Process Model for the Control of Information System Projects," *Proceedings of the Tenth International Conference on Information Systems*, New York, pp. 185–197, December 1989.
- [40] S. C. Ondash, S. Maloney, and J. Huerta, "Large Project Simulation: A Power Tool for Project Management Analysis," *Proceedings of the 1988 Winter Simulation Conference*, Los Angeles, pp. 231–239, January 1988.
- [41] A. J. Perlis, "Software Engineering Education," *Software Engineering Techniques: Report on a Conference Sponsored by the Nat Science Committee*, edited by J. N. Baxton and B. Randell, October 1969.
- [42] L. H. Putnam, "A General Empirical Solution to the Macro Software Sizing and Estimating Problem," *IEEE Transactions on Software Engineering*, vol. 4, pp. 345–361, July 1978.
- [43] L. H. Putnam and A. Fitzsimmons, "Estimating Software Costs, Part I," *Data-mation*, vol. 25, pp. 189–198, September 1979.
- [44] A. Radice, "Productivity Measures in Software," *The Economics of Information Processing Vol. 2: Operations, Programming, and Software Models*, edited by R. Goldberg and H. Lorin, New York: John Wiley and Sons, 1982.
- [45] C. V. Ramamoorthy, K. H. Kim, and W. T. Chen, "Optional Placement of Software Monitors Aiding Systematic Testing," *IEEE Transactions on Software Engineering*, vol. 1, pp. 46–58, 1975.
- [46] M. R. Rose and R. Harmsen, "Using Sensitivity Analysis to Simplify Ecosystem Models: A Case Study," *Simulation*, vol. 31, no. 1, pp. 15–24, 1978.
- [47] B. R. Schlender, "How to Break the Software Logjam," *Fortune*, vol. 212, pp. 100–112, September 25, 1989.
- [48] R. F. Scott and D. B. Simmons, "Predicting Programming Productivity—A Communications Model," *IEEE Transactions on Software Engineering*, vol. 1, pp. 230–238, December 1975.
- [49] R. C. Tausworthe, *Deep Space Network Software Cost Estimation Model*, JPL Publication 81-7, Jet Propulsion Laboratory, Pasadena, California, 1981.
- [50] R. C. Tausworthe, "Staffing Implication of Software Productivity Models," *TDA Progress Report 42-72*, vol. October–December, Jet Propulsion Laboratory, Pasadena, California, pp. 70–75, February 15, 1983.
- [51] T. A. Thayer, M. Lipow, and E. C. Nelson, *Software Reliability: A Study of Large Project Reality*, New York: North Holland, 1978.
- [52] R. W. Thayer, A. Pyster, and R. Wood, "Major Issues in Software Engineering Project Management," *IEEE Transactions on Software Engineering*, vol. SE-7, no. 4, pp. 111–125, July 1981.

- [53] R. D. H. Warburton, "Managing and Predicting the Costs of Real-Time Software," *IEEE Transactions in Software Engineering*, vol. 9, pp. 502–509, 1983.
- [54] L. G. Williams, "Software Process Modeling: A Behavioral Approach," *IEEE Transactions on Software Engineering*, vol. 4, pp. 174–186, 1978.
- [55] M. V. Zelkowitz, "Perspectives on Software Engineering," *Computing Surveys*, vol. 10, no. 2, pp. 197–216, June 1978.

Table 1. Staff members' evaluations of sensitivity-analysis tests.

Staff member		Overall evaluation of sensitivity test	
Number	Category	Reasonable = 1, Not reasonable = 0	Standard deviations
18	Manager	0.89	0.06
5	Researcher	0.87	0.18
23	Stratified	0.88	0.02

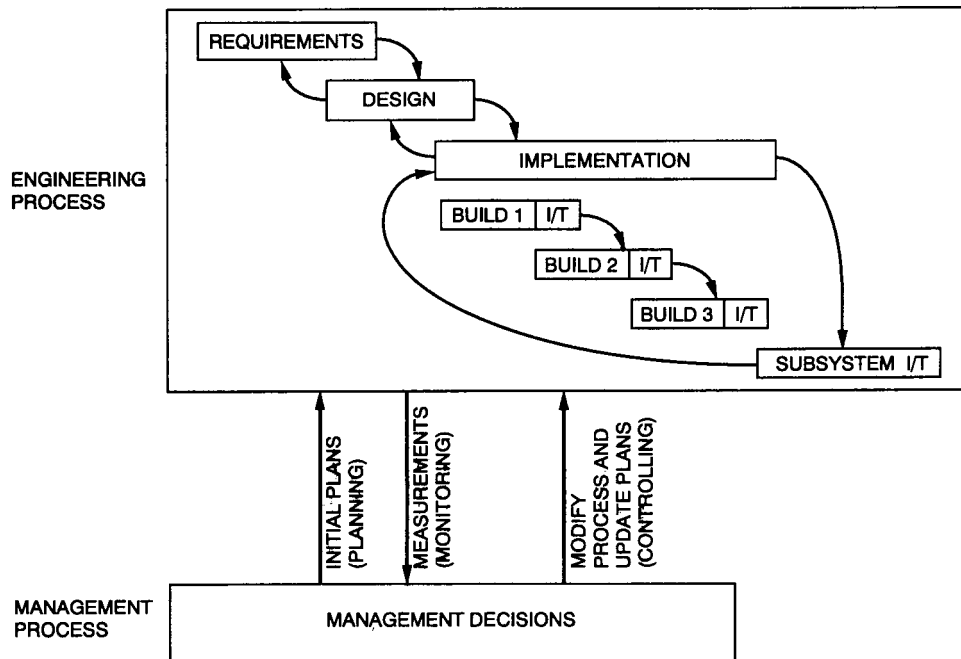


Fig. 1. Software engineering and management processes feedback structure.

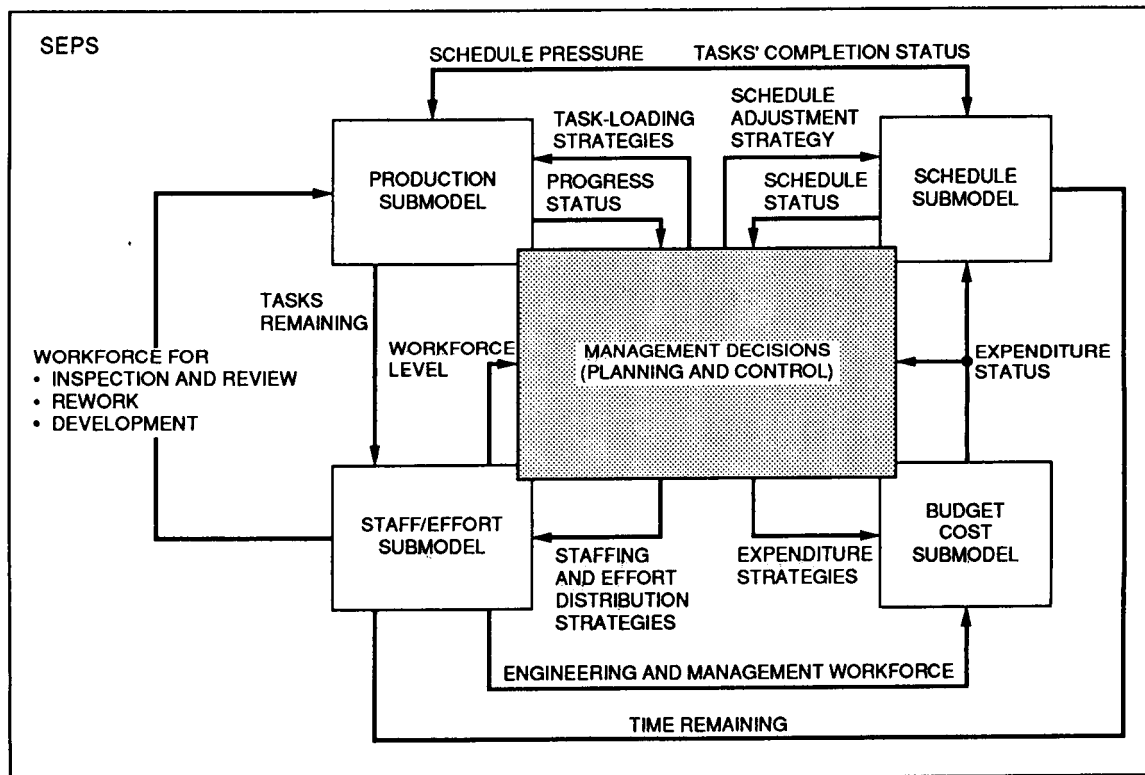


Fig. 2. Software project management model infrastructure.

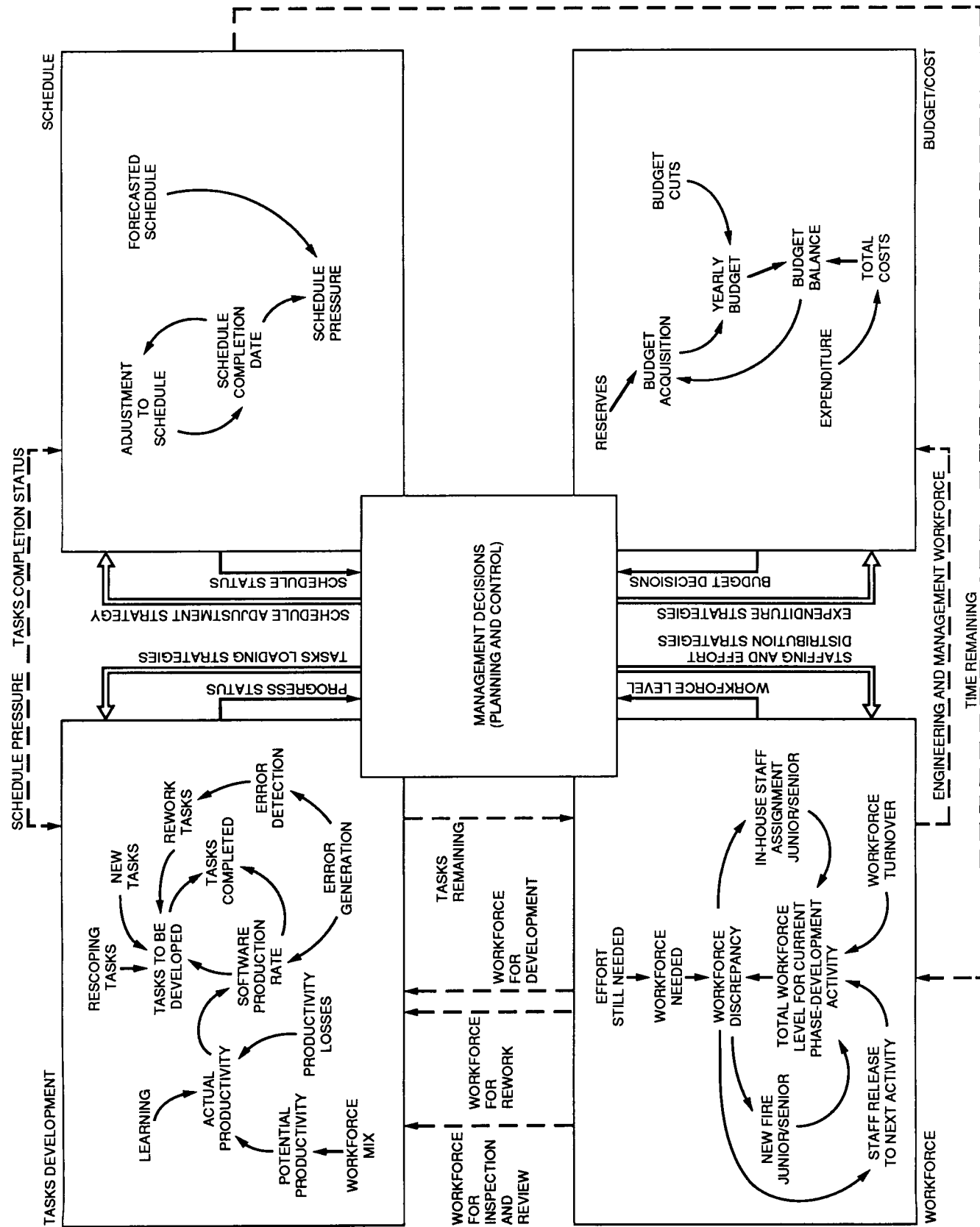


Fig. 3. Software engineering process feedback structure.

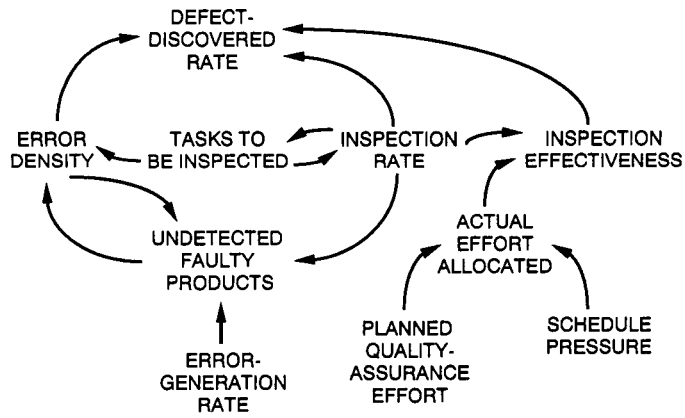


Fig. 4. Error-detection-rate submodel.



Fig. 6. Nominal errors generated per ksloc.

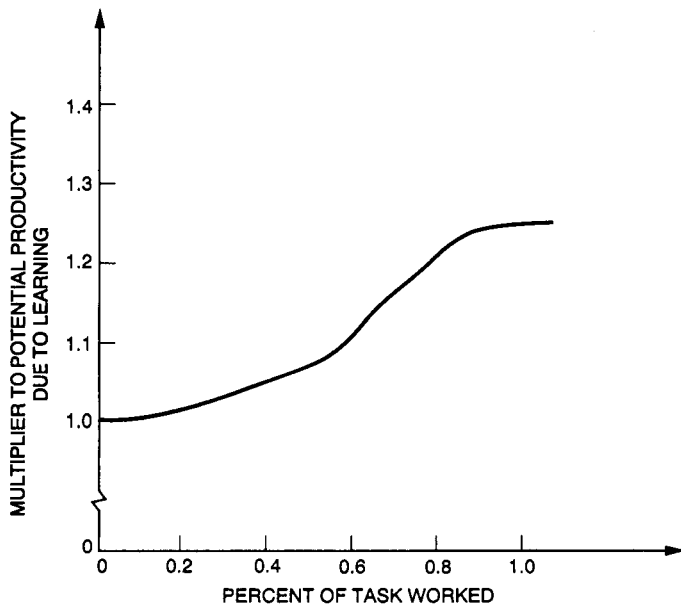


Fig. 5. Learning factor on potential productivity rate.

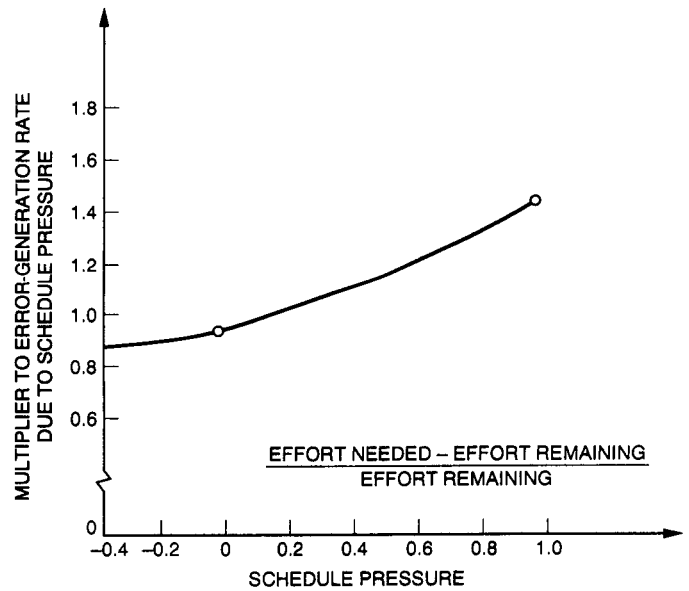


Fig. 7. Schedule pressure factor on error-generation rate.

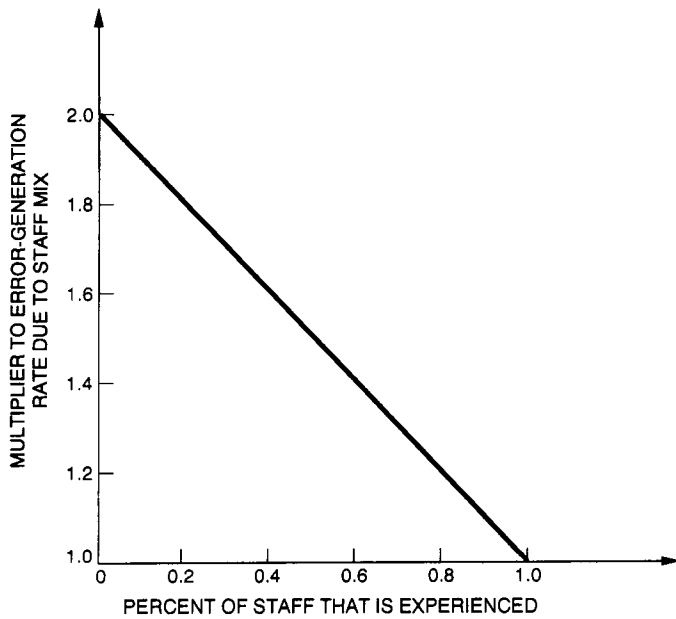


Fig. 8. Staff-mix effect on error-generation rate.

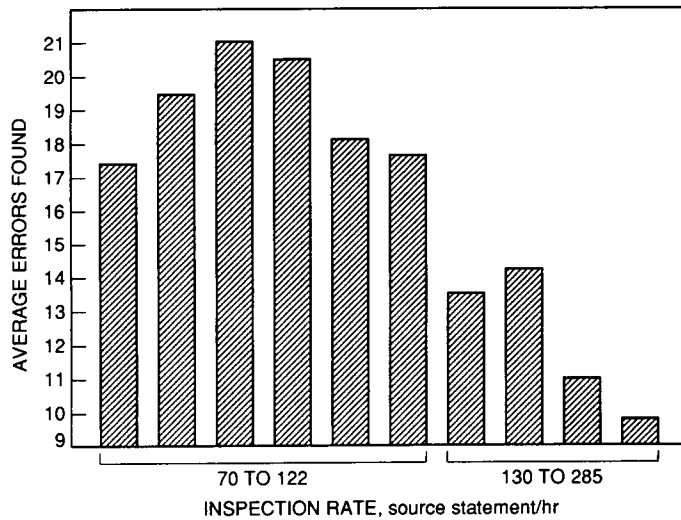


Fig. 9. Error-detection rate.

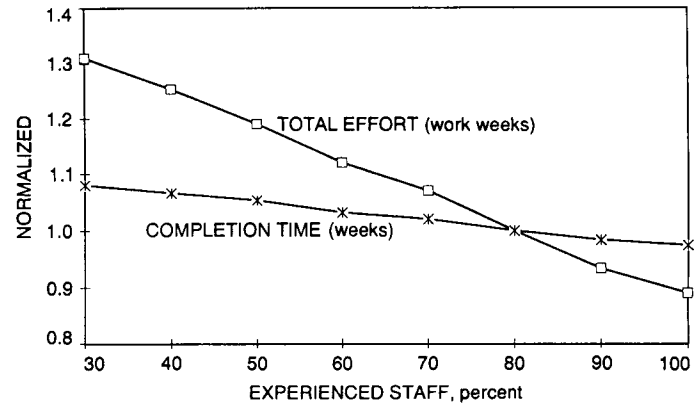


Fig. 10. The impact of staff experience level (percent) on project total effort and completion time normalized at 80-percent staff experience.

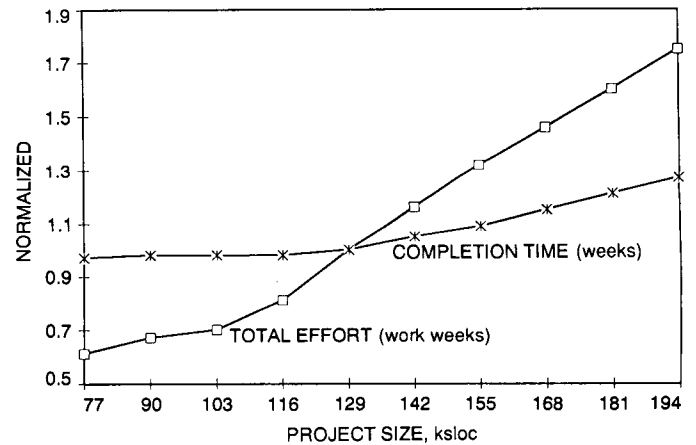


Fig. 11. The impact of project size (one thousand source lines of code) variations on project total effort and completion time normalized at 128 ksloc.

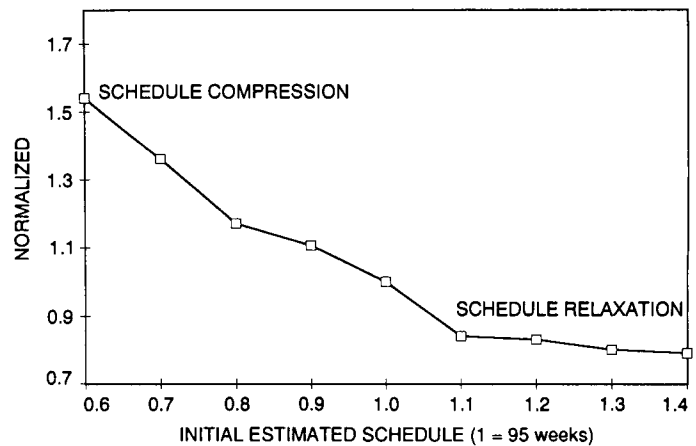


Fig. 12. The impact of schedule compression/relaxation on project total effort (work weeks), normalized at 1 = 95 weeks.

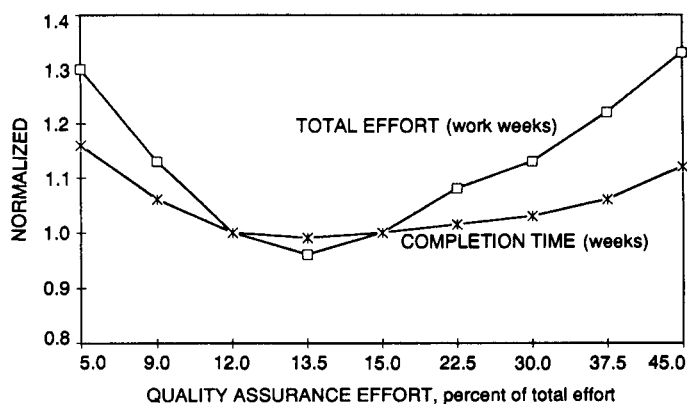


Fig. 13. The Impact of quality-assurance effort on project total effort and completion time normalized at QA effort of 15 percent of total effort.

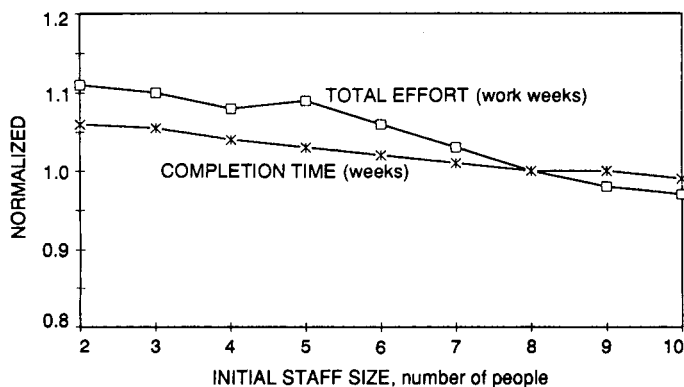


Fig. 14. The Impact of initial staff size variations on project total effort and completion time normalized at an initial staff size of 8 people.

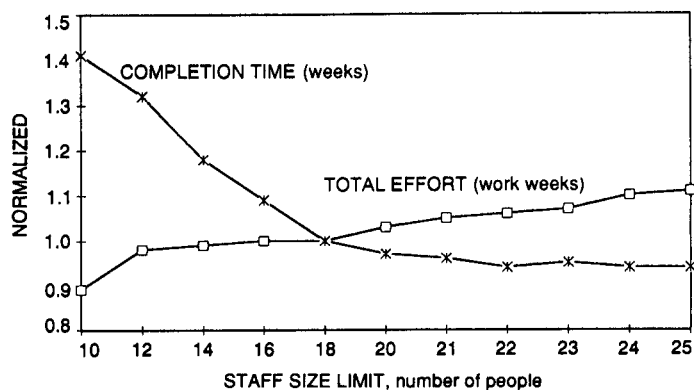


Fig. 15. The impact of staff size limit on project total effort and completion time normalized at a staff size limit of 18 people.

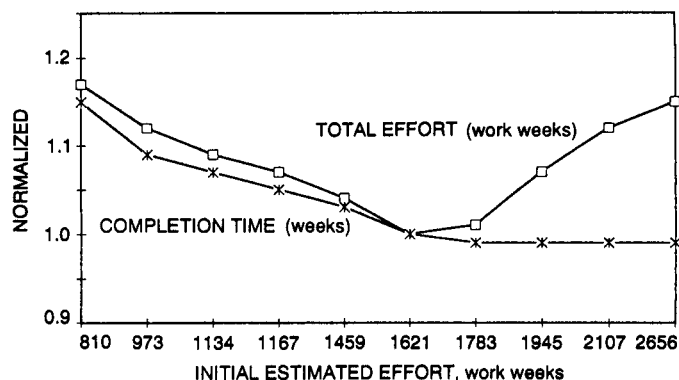


Fig. 16. The impact of initial estimated effort variation on project total effort and completion time normalized at initial estimated effort at 1,621 work weeks.

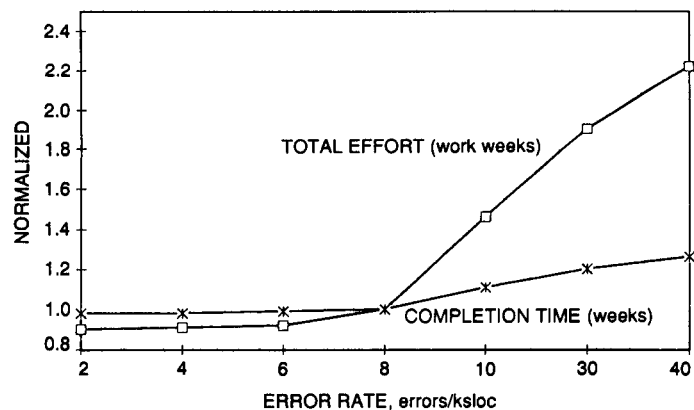


Fig. 17. The impact of error-rate variations on total effort and completion time normalized at 10 errors/ksloc.

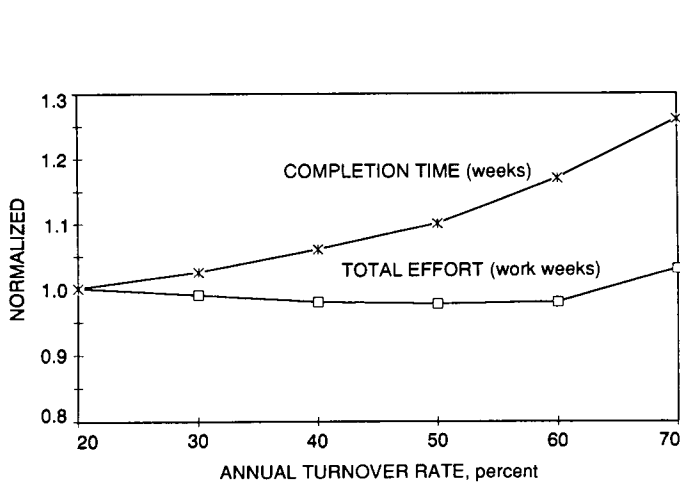


Fig. 18. The impact of annual turnover rate on project total effort and completion time normalized at annual turnover rate of 20 percent.

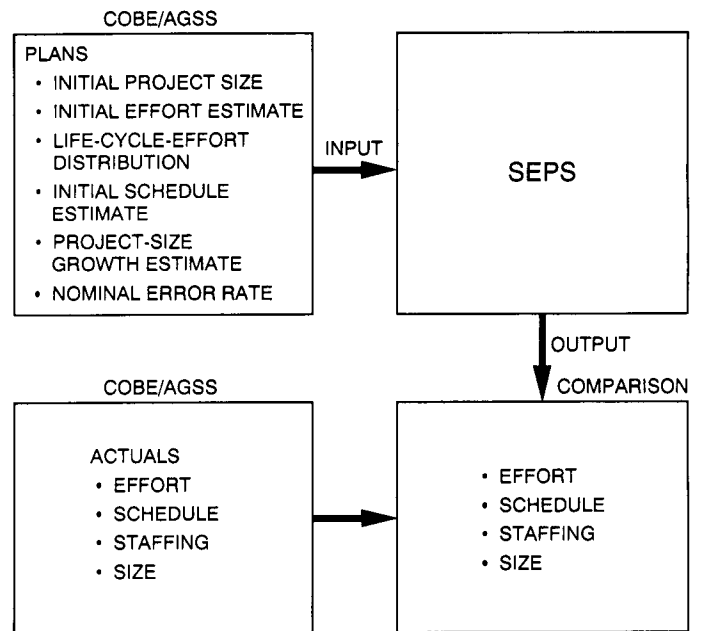


Fig. 19. Historical project-case comparison.

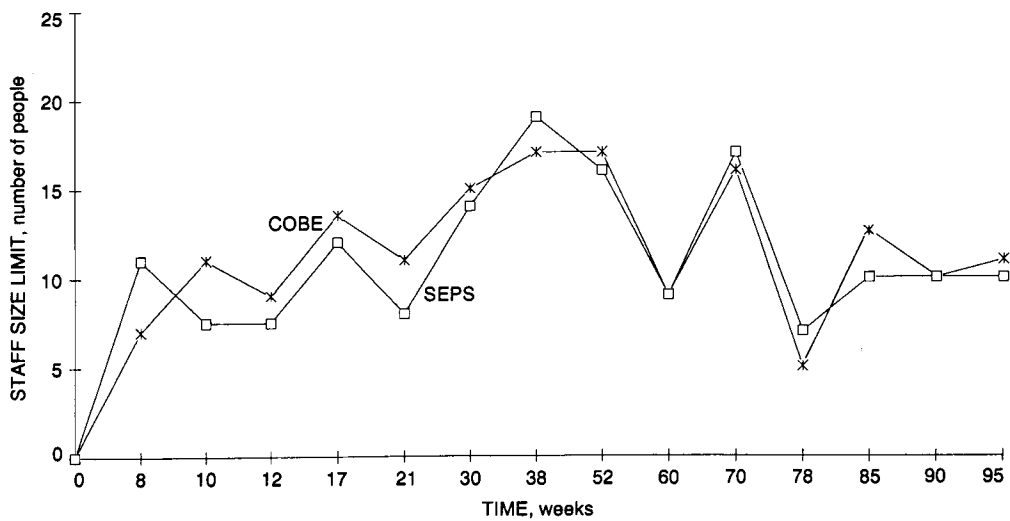


Fig. 20. Staffing-curve comparison.

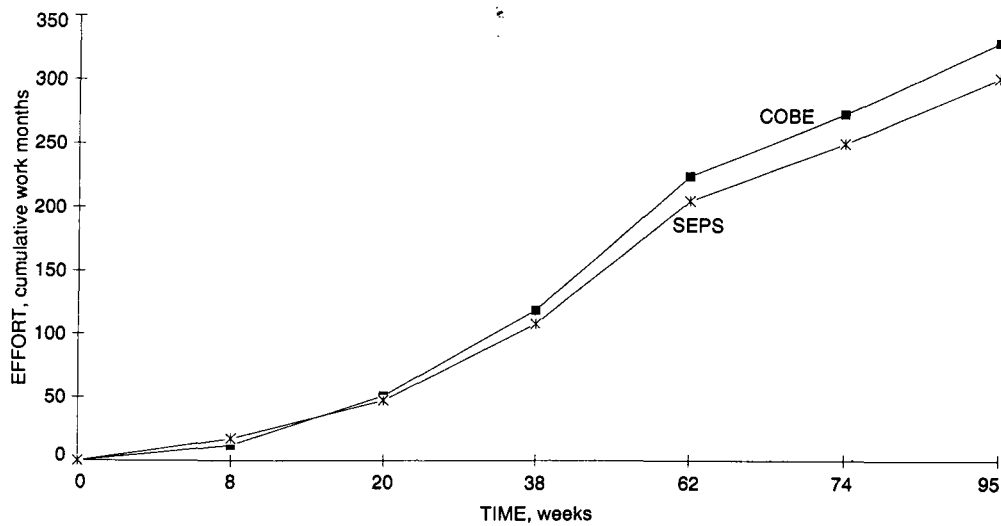


Fig. 21. Accumulated effort comparison of COBE and SEPS.

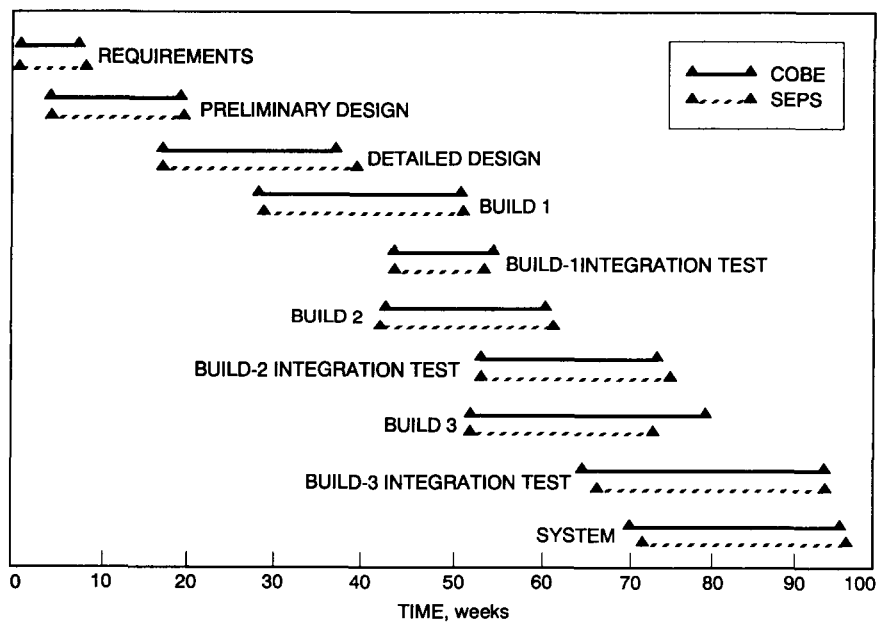


Fig. 22. Schedule comparison.

514-32

775-14

P-15

N92-24315

JJ570450

Node Synchronization Schemes for the Big Viterbi Decoder

K.-M. Cheung, L. Swanson, and S. Arnold
Communications Systems Research Section

The Big Viterbi Decoder (BVD), currently under development for the DSN, includes three separate algorithms to acquire and maintain node and frame synchronization. The first measures the number of decoded bits between two consecutive renormalization operations (renorm rate), the second detects the presence of the frame marker in the decoded bit stream (bit correlation), while the third searches for an encoded version of the frame marker in the encoded input stream (symbol correlation). This article gives a detailed account of the operation, as well as performance comparison, of the three methods.

I. Introduction

This article summarizes the node synchronization (abbreviated as node sync) study for the the Big Viterbi Decoder (BVD) currently under development for the DSN. The function of node sync is to align incoming channel symbols with the code trellis used by a Viterbi decoder and ensure proper phasing between the decoded bit clock and the received symbol clock. The proper phasing of the two clocks is crucial to the operation of the decoder and must be established before synchronized decoding starts, and monitored thereafter.

The present NASA standard concatenated code uses a (7,1/2) convolutional code as its inner code and an 8-bit (255,223) Reed-Solomon (RS) code as its outer code. This system achieves a bit error rate (BER) of 10^{-6} at a bit signal-to-noise ratio (SNR) of 2.53 dB. Recent code search efforts [1,2] show that SNR improvement of up to 2 dB is possible by using a constraint-length 15 convolutional code as the inner code. To demonstrate this performance improvement, an experimental (15,1/4) convolu-

tional encoder was implemented on the Galileo spacecraft, and a programmable convolutional decoder, the BVD hardware prototype, was completed and tested for codes with constraint lengths up to 15 and code rates of $1/N$, $N = 2, 3, \dots, 6$ [3]. Good node and frame synchronization schemes are essential to the realization of the aforementioned performance offered by long constraint length codes, which are expected to operate at very low SNR [the nominal operational SNR for the (15,1/4) convolutional code concatenated with the (255,223) RS code is $E_b/N_0 = 0.9$ dB, or equivalently $E_s/N_0 = -5.7$]. Since the outer RS code does not affect the node sync operations of the system, all subsequent discussions in this article address only node sync of the BVD, and SNR always means E_b/N_0 as seen by the BVD.

The BVD provides three separate algorithms to acquire and maintain node synchronization. Each of the three schemes operates in two modes: the acquisition mode, when the decoder is first activated or when a loss-of-sync is detected, and the tracking mode, when the decoder has

acquired sync and tries to maintain it. A false declaration of loss-of-sync during tracking, when the decoder is actually in-sync, is a severe offense that causes loss of valuable data and should be avoided. However, when the decoder truly loses sync, it should detect it and correct for it as soon as possible. A general DSN requirement during tracking is for no false loss-of-sync for 24 hr. For Galileo data rates, 134.4 kbits/sec and 115 kbits/sec, and a frame size of 5120 bits, the DSN requirement can be guaranteed by keeping the probability of false declaration of loss-of-sync below 4.4×10^{-7} .

Each of the three schemes is designed to individually meet the DSN requirements of node sync acquisition and tracking, so that the overall proposed system performance, which is based on a combination of the three schemes, is guaranteed to meet the DSN's requirements.

Scheme 1 measures the number of decoded bits between two consecutive renormalizations (renorm rate). It is a continuous operation and assumes no knowledge of the data format. This method can be used for node sync only. Schemes 2 and 3 require that an a priori known bit pattern, the frame marker, be inserted at regular intervals, known as frames, in the bit stream. Scheme 2 detects the presence of the frame marker in the decoded bit stream (bit correlation), while scheme 3 searches for an encoded version of the frame marker in the encoded input stream (symbol correlation). Schemes 2 and 3 are used for both node sync and frame sync. In the acquisition mode, a statistic x is measured at all possible locations of the frame marker, and the observed values of x are used to identify likely locations of the marker. Frame-to-frame verification (as in scheme 2) or integration over several frames (as in scheme 3) is used to reduce the probability of miss and the probability of false alarm at the expense of longer acquisition time. In the tracking mode the statistic x is monitored from frame to frame at the selected marker location to verify continued sync (flywheeling).

The three schemes have their respective advantages and disadvantages. The approach taken in the BVD is to use a combination of the three methods so that overall system performance is optimized. Previous research [4] indicates that each algorithm generates sync statistics that might be more reliable than others in a certain operational environment (e.g., SNR, frame size, operational mode, etc.). The overall node sync performance depends on the selection of the right combination of algorithms under the given operational environment.

The node sync analysis of the aforementioned three schemes involves many different user-chosen system pa-

rameters, such as code, frame size, frame-marker size, operating SNR, integration time, and so on. It is impossible to describe all the analysis results for many different combinations of parameters. In this article, the authors only present analysis results for the (15,1/4) code and the (15,1/6) code, with frame length $B = 8960$ bits, frame-marker size $L = 32$ bits, and acquisition time and tracking time (for schemes 2 and 3) equal to 4 frames. The authors have developed some software modules for the three schemes, which, when given the system parameters, generate the required thresholds and the corresponding synchronization performances.

By using the global node-sync procedure suggested in Section V, node sync will be acquired in less than 40,000 bits with a probability of 0.996, and true loss-of-sync will be detected in less than 40,000 bits with a probability of 0.99998.

The rest of the article is organized as follows: Section II describes the operations of scheme 1 and introduces a maximum-likelihood method to estimate the optimal threshold. Sections III and IV describe schemes 2 and 3 and their respective methods to determine the thresholds. Section V gives the performance comparisons of the three node sync schemes, Section VI discusses the robustness of the three schemes with respect to an uncertain SNR estimation, and finally Section VII discusses unresolved issues.

II. Renormalization Rate Scheme (Scheme 1)

This method measures the observable x , the number of bits between two consecutive state metric-renormalization events (renorms), or the mean number of bits between n consecutive renorms (the total number of bits between n consecutive renorms divided by n), and compares x with a preselected threshold. A similar scheme that measures the renorm counts within a long but fixed integration time was analyzed in [9] for the (7,1/2) code.

The renormalization-rate scheme starts with an arbitrary channel-symbol offset and tests the renormalization rate to detect node sync. If the system is determined to be out-of-sync, the offset is changed and the bit clock is changed by one channel symbol. This may need to occur as many as N times. Let H_0 be the hypothesis that the code is in-sync, and H_1 that the code is not. It is observed from the BVD software simulations that the conditional probability density function (pdf) $f(x|H_0)$ is approximately Gaussian with mean m_0 and variance σ_0^2 . Similarly, the conditional pdf $f(x|H_1)$ is approximately Gaussian with mean m_1 and variance σ_1^2 . That is,

$$f(x|H_0) \approx \frac{1}{\sqrt{2\pi}\sigma_0} \exp \frac{(x - m_0)^2}{2\sigma_0^2}$$

and

$$f(x|H_1) \approx \frac{1}{\sqrt{2\pi}\sigma_1} \exp \frac{(x - m_1)^2}{2\sigma_1^2}$$

These approximated pdf's are especially good when n is large (Central Limit Theorem). Then m_0 , m_1 , σ_0 , and σ_1 can be estimated through software simulation of the metric renorm of the BVD. By assuming $f(x|H_0)$ and $f(x|H_1)$ to be truly Gaussian, the probability of miss, P_M , which is the likelihood that the node sync scheme detects out-of-sync when the system is actually in-sync, is given by

$$P_M = P(x < T|H_0) = \int_{-\infty}^T f(x|H_0)dx$$

where the statistic x is compared with a threshold T . Similarly, the probability of false alarm, P_F , which is the likelihood that the node sync scheme detects in-sync when the system is actually out of sync, is given by

$$P_M = P(x > T|H_1) = \int_T^{\infty} f(x|H_1)dx$$

In the acquisition mode, node sync is achieved when the observable x is greater than or equal to the acquisition threshold T_1 at the correct location for which H_0 is true, and is smaller than T_1 at the $N - 1$ locations for which H_1 is true. By assuming the ideal autocorrelation property of the frame sync marker, one has

$$P_{acq} = (1 - P_F)^{N-1}(1 - P_M)$$

Tables 1 and 2 tabulate the optimal threshold and its corresponding P_{acq} for various SNR's for the (15,1/4) code and the (15,1/6) code, respectively. Notice that for a given E_b/N_0 , T_1 remains the same for different values of n , but P_{acq} falls off rapidly as n increases. The required acquisition time of this scheme, in the worst case, is nN renorm counts, since this scheme requires n renorms for each trial and there are at most N trials for a $1/N$ code. This corresponds to less than 40,000 bits in the worst case.

In the tracking mode, in order to maintain the false loss-of-sync probability $1 - P_{tk}$ below 4.4×10^{-7} , the threshold T_2 is chosen so that

$$1 - P_{tk} = P(x < T_2|H_0) \leq 4.4 \times 10^{-7}$$

The corresponding probability of detecting true loss-of-sync is then given by

$$P(x < T_2|H_1) = \int_{-\infty}^{T_2} f(x|H_1)dx$$

Tables 3 and 4 tabulate the threshold required to maintain the false loss-of-sync probability below 4.4×10^{-7} , and the corresponding probability of detecting true loss-of-sync for the (15,1/4) code and the (15,1/6) code, respectively.

III. Decoded Bit-Correlation Scheme (Scheme 2)

The decoded bit-correlation scheme is the DSN's current method for frame sync. This scheme is now available in the BVD to detect frame sync as well as node sync, because when frame sync is achieved, node sync naturally follows. The frame sync acquisition probabilities and their optimal thresholds for the (15,1/4) code at various SNR's were given in [8]. In this section, the authors evaluate the node sync acquisition and tracking probabilities and their corresponding optimal thresholds for the (15,1/4) code as well as the (15,1/6) code. The bit-correlation scheme alone is not an effective way for acquiring node sync. Node sync has to be achieved before the Viterbi decoder can operate properly, and this scheme operates on the decoded bit stream. To acquire node sync with this method, one needs to detect node sync by performing bit correlation on the B locations of a frame for as many as N possible channel symbol offsets. This may require as many as N frames to detect sync (and one more frame to verify sync), and such a long acquisition time does not meet DSN requirements. This scheme should only be used to verify schemes 1 and 3, that is, to check for the existence of the frame sync marker in the decoded bit stream, after node sync is established by scheme 1 or 3. The subsequent discussion assumes that correct node sync has already been acquired by schemes 1 and/or 3, and evaluates the frame sync acquisition performance of this scheme to see how well it can be used to verify schemes 1 and 3. In the tracking mode, this scheme gives a performance comparable to the other two schemes, and should be considered on an equal basis with the other two schemes.

Scheme 2 compares the true sync word marker to a 32-bit segment of decoded bits. Those bits found to be in disagreement are counted. This count is then compared with a predetermined threshold T_1 optimized for a given SNR. If the number of disagreements is greater than T_1 , those 32 bits are rejected as the sync word. Otherwise, the 32 bits are recorded as a sync word candidate. During the acquisition mode, successive one-bit shifts of 32-bit decoded signal segments are compared with the true sync word until the threshold test is passed at the same location in two consecutive frames. In this article, the authors wish to acquire sync with next-frame verification within four frames. Once sync has been declared, the scheme enters into the tracking mode, testing for the sync word is done in a “flywheeling” fashion, that is, by comparing the counts with a preset threshold T_2 only at the presumed frame-marker position. If two consecutive frames fail the threshold test, the scheme will abort the flywheeling operation and switch to the acquisition mode.

The decoded data bits were generated by using the Little Viterbi Decoder (LVD) [5], a hardware decoder that decodes a bit stream at a rate of 60 bps for the (15,1/4) convolutional code, and at a rate of 40 bps for the (15,1/6) code. The received symbols fed into the decoder represent encoded symbols generated from the all-zero information bit sequence corrupted by additive white Gaussian noise (AWGN). The LVD generated enough data to ensure that 100 error bursts were produced for each SNR of interest.

For each SNR tested, the decoded bits were subjected to the threshold test for possible threshold values T_1 , where $0 \leq T_1 \leq 10$. For a random 32-bit window of decoded bits, this test determines whether the number of decoded bit errors in the observed window exceeds the given threshold. A count is maintained of the number of 32-bit windows with more than T_1 errors. The 32-bit window is then shifted to the right by one bit until all possible 32-bit segments have been tested.

Due to the slow speed of the LVD, the authors only obtained enough data to do the sync analysis for the (15,1/4) code and the (15,1/6) code in the SNR range from 0 to 1 dB. Analysis in the higher SNR range probably requires the use of the BVD. This would involve modifying the software and hardware of the BVD, and that is something the authors would like to do in the near future.

In the decoded bit-correlation scheme, the acquisition probability P_{acq} and the false loss-of-sync probability $1 - P_{tk}$ in the tracking mode are functions of the probability of miss, P_M , and the probability of false alarm, P_F . P_M is the likelihood that the sync word is not detected in the decoded bit stream, which is $P(x \geq T | \text{sync marker in the current 32-bit window})$, where T is a preselected threshold. P_F is the likelihood that the sync word is falsely detected in an incorrect position in the decoded bit stream, which is $P(x < T | \text{sync marker not in the current 32-bit window})$. From the above discussion P_M can be estimated from the LVD error data as

$$P_M = \frac{\text{number of 32-bit windows where the number of errors exceeds } T}{\text{number of 32-bit windows tested within a given file}}$$

Assuming that random data P_F are given by [6],

$$P_F = \sum_{k=0}^T \binom{32}{k} 2^{-32}$$

Note that P_M depends on the code and the SNR, but P_F does not. Figures 1 and 2 give the P_M versus P_F curves for the (15,1/4) code and the (15,1/6) code, respectively.

In the acquisition mode, the probability of acquiring sync correctly with next-frame verification within four frames can be approximated for small values of P_M and P_F by [4]

$$P_{acq} \approx 1 - 3P_M^2 - \frac{B-1}{2}P_F^2$$

where $B = 8960$ is the length of a frame in bits. Tables 5 and 6 give the optimal threshold together with the corresponding P_{acq} for the (15,1/4) code and the (15,1/6) code, respectively. It is observed that in both cases the optimal threshold remains relatively constant for a wide range of SNR.

In the tracking mode, the node sync scheme is required to maintain the probability of false loss-of-sync declaration below 4.4×10^{-7} and to keep a reasonably high probability of true loss-of-sync declaration. On the basis of

their extensive simulation results, the authors propose to declare a loss-of-track when five consecutive frames fail the threshold test. To satisfy the stringent DSN requirement of keeping the false loss-of-sync probability below 4.4×10^{-7} , while at the same time keeping a high probability of detecting true loss-of-track, this is equivalent to finding the threshold T_2 so that

$$P_M^5 \leq 4.4 \times 10^{-7}$$

Tables 7 and 8 give the threshold values that achieve the above-required false loss-of-sync probability, together with the corresponding probability of detecting true loss-of-sync during a track, which is $(1 - P_F)^5$, for the (15,1/4) code and the (15,1/6) code, respectively.

IV. Symbol Correlation Scheme (Scheme 3)

The frame-sync acquisition probabilities (scheme 2) and their optimal thresholds for the (7,1/2) code at various SNR's were evaluated in [4]. In this section, the authors evaluate the node-sync acquisition and tracking probabilities and their corresponding optimal threshold of the (15,1/4) code as well as the (15,1/6) code. To describe the node-sync scheme based on channel-symbol measurements, the notation used in [4] is employed. Due to the low SNR operation range of the BVD, this node-sync scheme requires multiple-frame integration, and differs slightly from the frame-sync scheme proposed in [4]. To make this article more self-contained, the authors take into account the multiple-frame integration feature and rederive some of the equations needed for the node-sync analysis. The incoming data bit stream b_i includes both true data bits and sync marker bits λ_i . To simplify subsequent discussion, the encoded sync pattern is assumed to have the ideal autocorrelation property. The data bit stream is packaged into data frames b_i , $i = 1, \dots, B$, of B bits each, and L sync marker bits λ_i , $i = 1, \dots, L$, are included in every data frame. The data bit stream is convolutionally encoded by a rate $1/N$, constraint-length K convolutional encoder. The encoded channel symbol stream s_i , $i = 1, \dots, S$, is likewise partitioned into frames of $S = NB$ symbols each, and each frame includes a set of $M = N(L - K + 1)$ sync marker symbols m_i , $i = 1, \dots, M$, that are totally determined by the sync marker bits λ_i , $i = 1, \dots, L$. The remaining $N(B - L + K - 1)$ symbols in each frame depend solely on the true data bits or else on a combination of true data bits and sync marker bits.

The channel symbols are assumed to have constant magnitude s (i.e., $s_i = \pm s$), and they are received through

an AWGN channel with noise samples n_i , $i = 1, \dots, S$, with zero mean and variance σ^2 . The ratio $\rho = s^2/\sigma^2$ is an SNR parameter. In terms of ρ , the channel symbol SNR is $E_s/N_0 = \rho/2$, and the bit SNR is $E_b/N_0 = N\rho/2$. The received symbols r_i , $i = 1, \dots, S$, are passed through a maximum likelihood convolutional decoder (Viterbi decoder) to obtain the decoded bits d_i , $i = 1, \dots, B$.

Two factors contribute to the difficulties of synchronization using channel symbols. First, the long constraint-length, low-rate codes are designed to operate in a low SNR environment in which the channel symbols are severely corrupted by noise. Second, an inherent drawback of this scheme is that only $L - K + 1$ bits of frame marker are usable in the correlation. The encoded symbols corresponding to the other $K - 1$ bits depend on the previous contents of the encoder shift register. To recover enough SNR for the correlation, integration over j ($j \geq 1$) frames is needed at low SNR and/or when L is small. This is equivalent to increasing the SNR by a factor of j in a one-frame symbol correlation. The sync time, however, is also increased by a factor of j .

Let x be the symbol correlation statistic defined by¹

$$x = \sum_{i=1}^M m_i r_i$$

The BVD symbol-correlation scheme works as follows: When the decoder is first activated, or when a loss-of-sync is detected, the sync system initiates the acquisition mode. The statistic x is measured at all S possible locations of the frame marker, each integrated over j_1 frames, and the observed values of x are compared with a programmable threshold T_1 to identify likely locations of the marker. The statistic x is said to pass the threshold test if $x \geq T_1$. Sync is declared if only one value of x passes the threshold test and the rest fail. If all S values fail the threshold test, or when two or more values of x pass the threshold test, the decoder aborts the sync process and starts a new sync search.

After the decoder has acquired sync, the sync system initiates the tracking mode. The statistic x is tested against a preselected threshold T_2 only at the presumed marker position integrated over j_2 frames. Again, x is said to pass the threshold test if $x \geq T_2$. If x fails the

¹ This "positive correlation" statistic differs slightly from the "negative correlation" statistic used in [4], but it is equivalent in performance and it corresponds to the actual statistic measured by the BVD.

threshold test, the out-of-sync hypothesis is declared, and the sync system switches to acquisition mode.

The BVD actually has the capability of using two thresholds $T_{1,H}$ and $T_{1,L}$ (acquisition mode) or $T_{2,H}$ and $T_{2,L}$ (tracking mode), where $T_{1,L} = -T_{1,H}$ and $T_{2,L} = -T_{2,H}$. The lower threshold $T_{1,L}$ or $T_{2,L}$ is used to identify the sync marker in the case of bit inversion, which results from the operation of the telemetry receiver and cannot be easily overcome. In this article, only single thresholds T_1 (acquisition mode) and T_2 (tracking mode) in the normal non-bit-inversal case are considered. The results are easily extendable to the case of double thresholds.

The general performance expressions in this article are derived for arbitrary combinations of the parameters K , N , L , M , B , and S . The following is the performance analysis of this scheme.

Let x be integrated over j frames. The observed value of this statistic should be near jMs^2 if r_i , $i = 1, \dots, M$, contains the marker, and otherwise should be near 0. It is therefore natural to compare the observed values of x with a preselected threshold T to make tentative yes-no decisions about the location of the marker, according to whether x falls below or exceeds T . That is

$$x \begin{array}{c} \text{in-sync} \\ \geq \\ \text{out-of-sync} \end{array} T$$

If x falls below T when x is measured at the true position of the marker, the rule gives a tentative decision that misses the sync-marker location. Conversely, if x is at least T when x is not measured at the true marker position, then the decision causes a false detection of sync or false alarm. As in scheme 2, the effectiveness of the sync system can be measured by two competing measures: the probability of miss (P_M) and the probability of false alarm (P_F), which are both functions of T and are defined as

$$P_M = \text{Prob}[x \geq T | \text{sync marker in the current } M$$

– symbol window]

and

$$P_F = \text{Prob}[x < T | \text{sync marker not in the current } M$$

– symbol window]

By measuring x at the marker after integrating over j frames, the channel-symbol correlation statistic x can be expressed as

$$x = \sum_{i=1}^M m_i(jm_i + n_i^{(j)}) = \sum_{i=1}^M u_i$$

where $j m_i + n_i^{(j)}$ is the i th received symbol integrated over j frames, $n_i^{(j)}$ is $N(0, j\sigma^2)$, u_i is $N(js^2, js^2\sigma^2)$, and x/σ^2 is $N(jMs^2/\sigma^2, jMs^2/\sigma^2) = N(jM\rho, jM\rho)$. Thus, the miss probability P_M is calculated simply as

$$P_M(T, j\rho) = Q\left(\frac{jM\rho - \frac{T}{\sigma^2}}{\sqrt{jM\rho}}\right)$$

Away from the marker, the correlation statistic x is a sum of conditionally Gaussian random variables, some with zero mean and some with nonzero mean:

$$\begin{aligned} x &= \sum_{i=1}^{M-w} (js^2 + m_i n_i^{(j)}) + \sum_{i=1}^w (-js^2 + m_i n_i^{(j)}) \\ &= \sum_{i=1}^{M-w} u_i + \sum_{i=1}^w v_i \end{aligned}$$

where $i = 1, \dots, w$ represents the indices of the encoded symbols that differ from the marker symbols; v_i is $N(-js^2, js^2\sigma^2)$; and x/σ^2 is $N(j(M-2w)s^2/\sigma^2, jMs^2/\sigma^2) = N(j(M-2w)\rho, jM\rho)$. The false alarm probability is obtained by averaging the conditional Gaussian probability distribution for x over the discrepancy weight distribution $\text{Prob}[w]$,

$$P_F(T, j\rho) = \sum_{w=0}^M \text{Prob}[w] Q\left(\frac{2wj\rho - jM\rho + \frac{T}{\sigma^2}}{\sqrt{jM\rho}}\right)$$

where $\text{Prob}[w] \approx 2^{-M} \binom{M}{w}$ [7]. Note that both $P_M(T, j\rho)$ and $P_F(T, j\rho)$ depend on the symbol SNR ρ and the integration interval j only in terms of the product $j\rho$, which is the effective symbol SNR after integration over j frames.

Figures 3 and 4 show $P_M(T, j\rho)$ versus $P_F(T, j\rho)$ for the symbol-correlation statistic (frame-marker length $L = 32$ bits and integration interval $j = 4$) for the (15,1/4) code and the (15,1/6) code, respectively. Note that doubling the

integration time is equivalent to increasing the SNR by a factor of 2. Thus, for example, integrating over 4 frames at 0.0 dB has the same performance as integrating over 2 frames at 3.0 dB.

In the acquisition mode, the statistic x is checked at all possible locations of the frame sync marker after integration over j_1 frames. The observed values of x are compared with a preset threshold T_1 . Sync is declared if only one of S values of x passes the threshold test ($x \geq T$) and the rest fail ($x < T$). With the assumption that the sync marker sequence has an ideal autocorrelation property, the statistics of x are independent from location to location, and P_M and P_F , as functions of threshold T_1 and integration time j_1 frames, determine the probability of acquisition P_{acq} integrated over j_1 frames as follows:

$$P_{\text{acq}}(T_1, j_1\rho) = (1 - P_F(T_1, j_1\rho))^{S-1} (1 - P_M(T_1, j_1\rho))$$

An optimal normalized threshold $M\rho - T_1/j_1\sigma^2$ that maximizes $P_{\text{acq}}(T_1, j_1\rho)$ can be evaluated for any given combination of SNR, code rate, frame length, frame-marker size, and integration time. Tables 9 and 10 give the optimal normalized thresholds $M\rho - T_1/j_1\sigma^2$ together with their corresponding P_{acq} with $j_1 = 4$ and $B = 8960$ for the (15,1/4) code and the (15,1/6) code, respectively.

In the tracking mode, the statistic x is integrated over j_2 frames, and is then compared with a preset threshold T_2 only at the presumed frame-marker position (fly-wheeling). If $x > T_2$, the in-sync hypothesis is assumed, otherwise the decoder declares out-of-sync. Tables 11 and 12 give the normalized thresholds $M\rho - T_2/j_2\sigma^2$ ($j_2 = 4$) and their corresponding probability of detecting true loss-of-sync during a track $1 - P_F(T_2, j_2\rho)$ for the (15,1/4) code and the (15,1/6) code, respectively, that produce a false loss-of-sync probability $P_M(T_2, j_2\rho)$ of 4.4×10^{-7} , with a respective frame marker length of 32 bits and frame length of 8960 bits.

V. Comparisons of the Three Schemes

The proposed BVD approach on node sync is to use a combination of the three methods so that the overall system performance is optimized. This section describes the advantages and disadvantages of the three node sync schemes. This section also suggests a global node sync procedure, which is based on the three node-sync algorithms described above, to declare in-sync and loss-of-sync, both in the acquisition mode and in the tracking mode.

From this study, scheme 1 gives the best node sync performance, both in the acquisition mode and tracking mode. However, a poor estimate of E_b/N_0 can degrade its performance quite rapidly. Scheme 2 alone, as mentioned in Section III, is not an effective method to acquire node-sync due to its long acquisition time. It is, however, more robust than schemes 1 and 3 and can be used to verify and confirm the other two schemes. In fact, when this scheme is used to verify node sync, it performs better than scheme 3 for $E_b/N_0 \geq 0.4$ dB. Scheme 3 gives a reasonably good node-sync performance, though not as good as scheme 1. However, scheme 3 is more robust than scheme 1 in the presence of poor E_b/N_0 estimation. Also, scheme 2 has the extra capability to detect symbol inversion, and it is the only scheme that can perform node sync without convolutional decoding.

Based on the above analysis, the authors propose the following global node sync procedure: In the acquisition mode, node sync is acquired when any two schemes declare sync. That is, when any two schemes pass the threshold tests (threshold values are given in Tables 1, 2, 5, 6, 9, and 10). In the tracking mode, a loss of track is declared when two out of three schemes fail the threshold tests (threshold values are given in Tables 3, 4, 7, 8, 11, and 12). This node sync procedure, which is obtained through computer simulations and modeling analysis, can be refined and improved when the node sync experiments are performed using the actual BVD hardware.

VI. Robustness of the Three Schemes With Respect to SNR

The three node sync schemes discussed above share one common feature: They all measure an observable x and compare it with a threshold T , which is a function of the operating SNR. Knowledge of the operating SNR is required for the three schemes to work properly. A good estimate of SNR is difficult to obtain, especially at a low SNR. However, the authors observe that in the above node sync schemes, either the threshold values change very little for a fairly wide range of SNR (schemes 2 and 3), or the node sync schemes operate so well that any threshold values corresponding to a ± 1 -dB range of SNR about the operating SNR can provide a synchronization performance that satisfies the DSN requirements (schemes 1 and 3) of fast node sync acquisition and very low false declaration of sync probability. Thus, an SNR estimation with ± 1 -dB accuracy is sufficient to guarantee the node sync schemes of the BVD to work properly. However, if the SNR estimation is off by more than 1 dB, the node-sync performance can degrade rapidly. Figures 5 and 7 give the

(15,1/4) code-acquisition performance of schemes 1 and 3, respectively, by using the threshold value at 0.5 dB when the true SNR ranges from 0–1.5 dB. Figure 6 gives the (15,1/4) code-acquisition performance of scheme 2 by using the 0.5-dB threshold value when the true SNR ranges from 0–0.8 dB.

VII. Unresolved Issues

- (1) Though accurate SNR estimation is not needed to ensure proper node synchronization of the BVD, techniques for SNR estimation, particularly at low SNR operation range, deserve more study as they are required in such other aspects of the BVD operation as choosing the proper stepsize for channel symbol quantization.
- (2) In all three schemes, it is possible to avoid SNR estimation and at the same time preserve a good sync acquisition performance by using the best-match approach instead of the threshold approach. The threshold approach, and thus SNR estimation, is still required in the tracking mode. The best-match approach requires storing all the observables that correspond to different synchronization offsets and picking the sync location that gives the most favorable observable. This increases the hardware and software complexity to various degrees (with respect to the threshold implementation), depending on the synchronization schemes. The best-match strategy is, in fact, equivalent to maximum-likelihood synchronization estimation, and can be shown to give better performance than the threshold strategy. Mathematical analysis and performance simulation of the three schemes using the best-match strategy are tedious and will not be given in this article.

The issue of implementation complexity is discussed as follows: Scheme 1 has only N different synchronization offsets and requires storing N values and choosing among N values. The increase in complexity is minimal. In fact, the current BVD prototype uses the best-match strategy in software to perform initial node sync acquisition in testing. However, best-match strategy for schemes 2 and 3 is not available in the BVD. Schemes 2 and 3 perform node sync as well as frame sync, and the use of the best-match strategy can enhance both node sync and frame sync. Since there are B possible sync offset locations in the decoded bit stream, scheme 2 requires a buffer of size B . Also scheme 2 requires the ability to pick in real time the most favorable observable among an array of B values. These functions do not yet exist in the current BVD frame-synchronization

subsystem. Scheme 3 is similar to scheme 2. It requires a buffer of size NB . However, since scheme 3 allows channel symbol integration over a few frames, this buffer exists in the current frame-synchronization subsystem and can be tapped into to perform the best-match function. Thus, the three sync schemes can be modified as suggested above to avoid SNR estimation and at the same time improve acquisition performance, since the best-match approach is in fact a maximum-likelihood strategy.

- (3) The current DSN method for frame sync, which is a bit-correlation scheme, uses a “test and verify” strategy for frame sync acquisition. That is, the unique sync candidate from any frame is chosen as the sync location if and only if it is verified once in the next succeeding frame by repassing the threshold test at the corresponding location within that frame. This scheme requires making two separate and independent frame sync decisions, one frame after the other. Intuitively, this scheme is inferior to the scheme that takes into account the bit-correlation information of two successive frames before making a decision on frame sync acquisition. One can call this scheme an “integrate and test” strategy. This strategy, in a rough sense, is equivalent to performing decoded bit correlation by using a frame marker of size $2L$ over a frame length of size $2B$. Table 13 tabulates the optimal threshold and its corresponding P_{acq} for various SNR for the (15,1/4) code by using the new decoded bit-correlation scheme. Figure 8 compares the P_{acq} performances between the old “test and verify” strategy and the new “integrate and test” strategy. The new strategy is uniformly better than the old scheme by about a factor of 2 on the probability of not acquiring sync within four frames. The additional hardware required is a buffer of size B , as there are B possible sync locations and the scheme requires storing a temporary correlation observable for each location in every other frame. On the other hand, this scheme does not require the logic to perform frame-to-frame verifications.
- (4) If SNR estimation cannot be achieved to within ± 1 -dB accuracy, either more sophisticated node-sync schemes [as suggested in issue (2)] should be used, or a 64-bit frame marker should be used. Figure 9 gives the (15,1/4) code-acquisition performance of scheme 2 using an 0.8-dB threshold value and a 64-bit sync marker when the true SNR ranges from 0–0.8 dB, and Fig. 10 gives the (15,1/4) code-acquisition performance of scheme 3 by using a 2-dB threshold value and a 64-bit sync marker when the true SNR ranges from 0–4 dB.

References

- [1] J. H. Yuen and Q. D. Vo, "In Search of a 2-dB Coding Gain," *TDA Progress Report 42-83*, vol. July–September 1985, Jet Propulsion Laboratory, Pasadena, California, pp. 26–33, November 15, 1985.
- [2] S. Dolinar, "A New Code for Galileo," *TDA Progress Report 42-93*, vol. January–March 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 83–96, May 15, 1988.
- [3] J. Statman, G. Zimmerman, F. Pollara, and O. Collins, "A Long Constraint Length VLSI Viterbi Decoder for the DSN," *TDA Progress Report 42-95*, vol. July–September 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 134–142, November 15, 1988.
- [4] S. Dolinar and K. Cheung, "Frame Synchronization Methods Based on Channel Symbol Measurements," *TDA Progress Report 42-98*, vol. April–June 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 121–137, August 15, 1989.
- [5] C. Lahmeyer and K. Cheung, "Long Decoding Runs for Galileo's Convolutional Codes," *TDA Progress Report 42-95*, vol. July–September 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 143–146, November 15, 1988.
- [6] L. Swanson, *A Comparison of Frame Synchronization Methods*, JPL Publication 82-100, Jet Propulsion Laboratory, Pasadena, California, December 15, 1982.
- [7] K. Cheung, "The Weight Distribution and Randomness of Linear Codes," *TDA Progress Report 42-97*, vol. January–March 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 208–215, August 15, 1989.
- [8] S. Arnold and L. Swanson, "Frame Synchronization for the Galileo Code," *TDA Progress Report 42-104*, vol. October–December 1990, Jet Propulsion Laboratory, Pasadena, California, pp. 211–218, February 15, 1991.
- [9] U. Cheng, "Node Synchronization of Viterbi Decoders Using State Metrics," *TDA Progress Report 42-94*, vol. April–June 1988, Jet Propulsion Laboratory, Pasadena, California, pp. 201–209, August 15, 1988.

Table 1. (15,1/4) code-acquisition performance of scheme 1.

SNR, dB	Renorm counts, n	Threshold, T_1	$1 - P_{acq}$
0.1	1	1749	7.69×10^{-10}
0.1	2	1746	$< 1.00 \times 10^{-16}$
0.5	1	1814	7.99×10^{-15}
0.5	2	1806	$< 1.00 \times 10^{-16}$
1.0	1	1901	$< 1.00 \times 10^{-16}$
1.0	2	1894	$< 1.00 \times 10^{-16}$
1.5	1	2001	$< 1.00 \times 10^{-16}$
1.5	2	1985	$< 1.00 \times 10^{-16}$
2.0	1	2106	$< 1.00 \times 10^{-16}$
2.0	2	2100	$< 1.00 \times 10^{-16}$
2.5	1	2221	$< 1.00 \times 10^{-16}$
2.5	2	2184	$< 1.00 \times 10^{-16}$
3.0	1	2305	$< 1.00 \times 10^{-16}$
3.0	2	2275	$< 1.00 \times 10^{-16}$

Table 3. (15,1/4) code-tracking performance of scheme 1.

SNR, dB	Renorm counts, n	Threshold, T_2	$1 - P$ (detect true loss-of-track)
0.1	1	1812	$< 1.00 \times 10^{-16}$
0.1	2	1886	$< 1.00 \times 10^{-16}$
0.5	1	1995	$< 1.00 \times 10^{-16}$
0.5	2	2087	$< 1.00 \times 10^{-16}$
1.0	1	2172	$< 1.00 \times 10^{-16}$
1.0	2	2256	$< 1.00 \times 10^{-16}$
1.5	1	2527	$< 1.00 \times 10^{-16}$
1.5	2	2616	$< 1.00 \times 10^{-16}$
2.0	1	2927	$< 1.00 \times 10^{-16}$
2.0	2	3040	$< 1.00 \times 10^{-16}$
2.5	1	3456	$< 1.00 \times 10^{-16}$
2.5	2	3558	$< 1.00 \times 10^{-16}$
3.0	1	4047	$< 1.00 \times 10^{-16}$
3.0	2	4173	$< 1.00 \times 10^{-16}$

Table 2. (15,1/6) code-acquisition performance of scheme 1.

SNR, dB	Renorm counts, n	Threshold, T_1	$1 - P_{acq}$
0.0	1	756	5.44×10^{-4}
0.0	2	754	7.30×10^{-7}
0.5	1	797	7.26×10^{-7}
0.5	2	792	5.09×10^{-13}
1.0	1	855	6.26×10^{-12}
1.0	2	849	$< 1.00 \times 10^{-16}$
1.5	1	886	$< 1.00 \times 10^{-16}$
1.5	2	882	$< 1.00 \times 10^{-16}$
2.0	1	952	$< 1.00 \times 10^{-16}$
2.0	2	949	$< 1.00 \times 10^{-16}$
2.5	1	965	$< 1.00 \times 10^{-16}$
2.5	2	960	$< 1.00 \times 10^{-16}$
3.0	1	1065	$< 1.00 \times 10^{-16}$
3.0	2	1054	$< 1.00 \times 10^{-16}$

Table 4. (15,1/6) code-tracking performance of scheme 1.

SNR, dB	Renorm counts, n	Threshold, T_2	$1 - P$ (detect true loss-of-track)
0.0	1	717	3.88×10^{-1}
0.0	2	753	8.25×10^{-8}
0.5	1	795	1.10×10^{-7}
0.5	2	838	$< 1.00 \times 10^{-16}$
1.0	1	906	$< 1.00 \times 10^{-16}$
1.0	2	953	$< 1.00 \times 10^{-16}$
1.5	1	1020	$< 1.00 \times 10^{-16}$
1.5	2	1078	$< 1.00 \times 10^{-16}$
2.0	1	1172	$< 1.00 \times 10^{-16}$
2.0	2	1225	$< 1.00 \times 10^{-16}$
2.5	1	1334	$< 1.00 \times 10^{-16}$
2.5	2	1407	$< 1.00 \times 10^{-16}$
3.0	1	1548	$< 1.00 \times 10^{-16}$
3.0	2	1619	$< 1.00 \times 10^{-16}$

Table 5. (15,1/4) code-acquisition performance of scheme 2.

SNR, dB	Threshold, T_1	$1 - P_{acq}$
0.0	6	8.59×10^{-3}
0.1	6	5.97×10^{-3}
0.2	6	3.29×10^{-3}
0.3	5	1.44×10^{-3}
0.4	5	7.00×10^{-4}
0.5	5	4.60×10^{-4}
0.6	5	2.60×10^{-4}
0.7	5	1.10×10^{-4}
0.8	4	5.00×10^{-5}

Table 6. (15,1/6) code-acquisition performance of scheme 2.

SNR, dB	Threshold, T_1	$1 - P_{acq}$
0.1	5	1.79×10^{-3}
0.3	5	4.94×10^{-4}
0.5	5	1.27×10^{-4}
0.7	4	3.35×10^{-5}
0.9	4	7.90×10^{-6}
1.1	3	1.70×10^{-6}

Table 7. (15,1/4) code-tracking performance of scheme 2.

SNR, dB	Threshold, T_2	$1 - P$ (detect true loss-of-track)
0.0	8	2.77×10^{-2}
0.1	7	8.38×10^{-3}
0.2	7	8.38×10^{-3}
0.3	7	8.38×10^{-3}
0.4	7	8.38×10^{-3}
0.5	7	8.38×10^{-3}
0.6	7	8.38×10^{-3}
0.7	7	8.38×10^{-3}
0.8	7	8.38×10^{-3}

Table 8. (15,1/6) code-tracking performance of scheme 2.

SNR, dB	Threshold, T_2	$1 - P$ (detect true loss-of-track)
0.1	7	8.38×10^{-3}
0.3	6	2.14×10^{-3}
0.5	6	2.14×10^{-3}
0.7	6	2.14×10^{-3}
0.9	6	2.14×10^{-3}
1.1	6	2.14×10^{-3}

Table 9. (15,1/4) code-acquisition performance of scheme 3.

SNR, dB	Threshold, T_1	$1 - P_{acq}$
0.0	8.7	3.69×10^{-3}
0.5	9.8	2.06×10^{-3}
1.0	11.0	1.12×10^{-3}
1.5	12.3	6.01×10^{-4}
2.0	13.7	3.15×10^{-4}
2.5	15.2	1.62×10^{-4}
3.0	16.9	8.21×10^{-5}

Table 10. (15,1/6) code-acquisition performance of scheme 3.

SNR, dB	Threshold, T_1	$1 - P_{acq}$
0.0	10.1	7.14×10^{-4}
0.5	11.4	3.21×10^{-4}
1.0	12.8	1.38×10^{-4}
1.5	14.4	5.74×10^{-5}
2.0	16.1	2.30×10^{-5}
2.5	17.9	8.90×10^{-6}
3.0	19.8	3.34×10^{-6}

Table 11. (15,1/4) code-tracking performance of scheme 3.

SNR, dB	Threshold, T_2	$1 - P$ (detect true loss-of-track)
0.0	14.8	1.90×10^{-5}
0.5	15.6	6.32×10^{-6}
1.0	16.6	2.04×10^{-6}
1.5	17.5	6.50×10^{-7}
2.0	18.6	2.00×10^{-7}
2.5	19.7	6.00×10^{-8}
3.0	20.8	1.00×10^{-8}

Table 12. (15,1/6) code-tracking performance of scheme 3.

SNR, dB	Threshold, T_2	$1 - P$ (detect true loss-of-track)
0.0	14.8	1.59×10^{-6}
0.5	15.6	3.50×10^{-7}
1.0	16.6	7.00×10^{-8}
1.5	17.5	1.00×10^{-9}
2.0	18.6	$< 1.00 \times 10^{-16}$
2.5	19.7	$< 1.00 \times 10^{-16}$
3.0	20.8	$< 1.00 \times 10^{-16}$

Table 13. (15,1/4) code-acquisition performance of the new scheme.

SNR, dB	Threshold, T_1	$1 - P_{acq}$
0.0	13	3.97×10^{-3}
0.2	13	1.28×10^{-3}
0.4	12	3.29×10^{-4}
0.6	12	7.55×10^{-5}
0.7	12	3.69×10^{-5}
0.8	11	1.40×10^{-5}

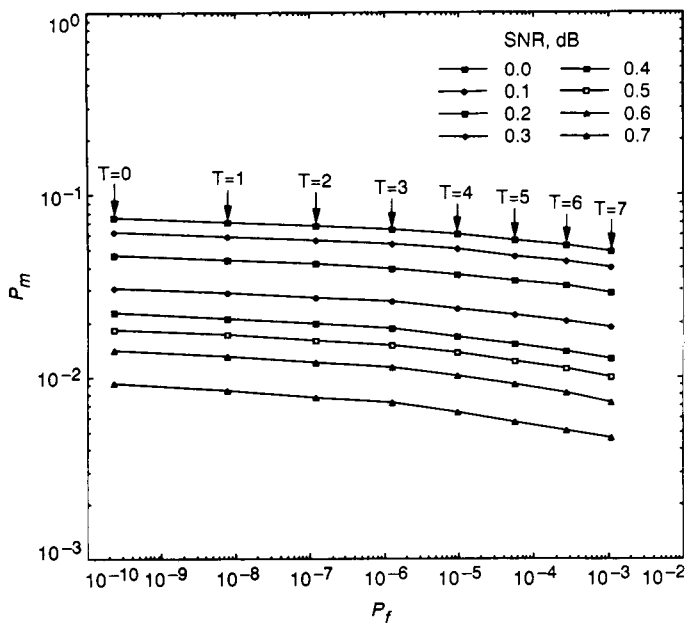


Fig. 1. P_m versus P_f (frame length = 8960 bits).

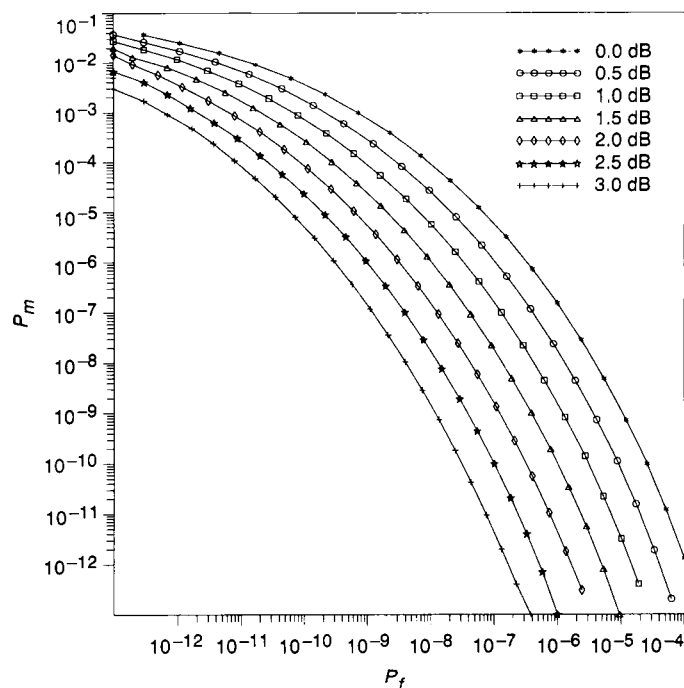


Fig. 3. P_m versus P_f for the (15,1/4) code ($J = 4$).

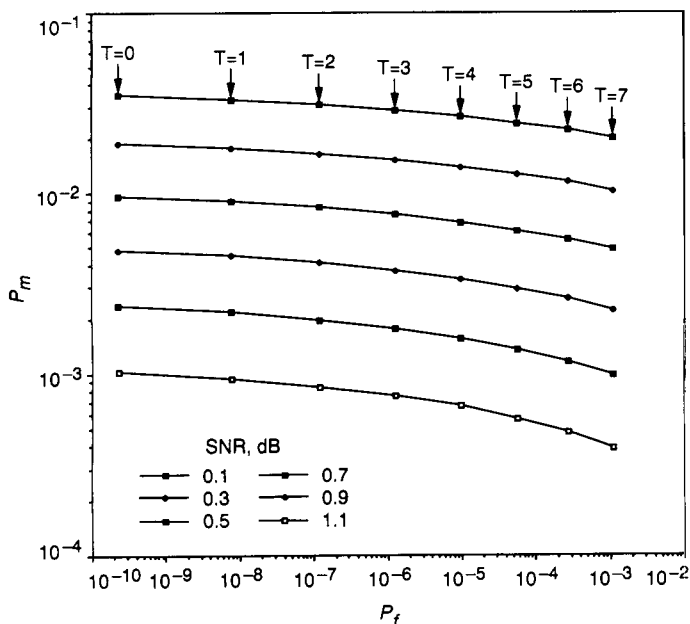


Fig. 2. P_m versus P_f (frame length = 8960 bits).

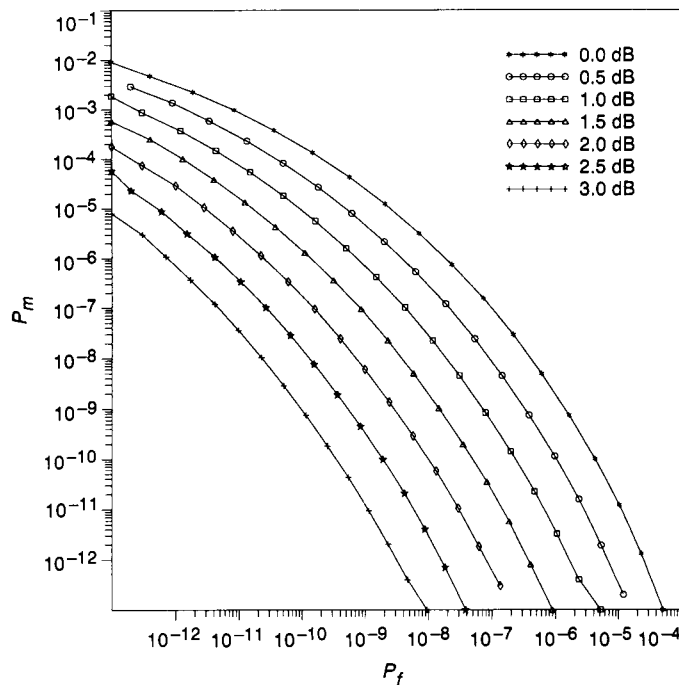


Fig. 4. P_m versus P_f for the (15,1/6) code ($J = 4$).

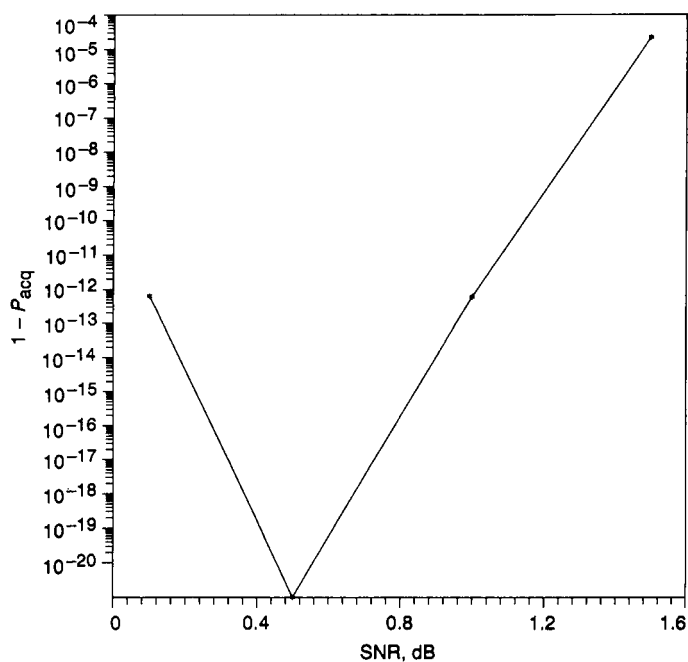


Fig. 5. (15,1/4) code-acquisition performance of scheme 1 using a fixed threshold from 0–1.5 dB.

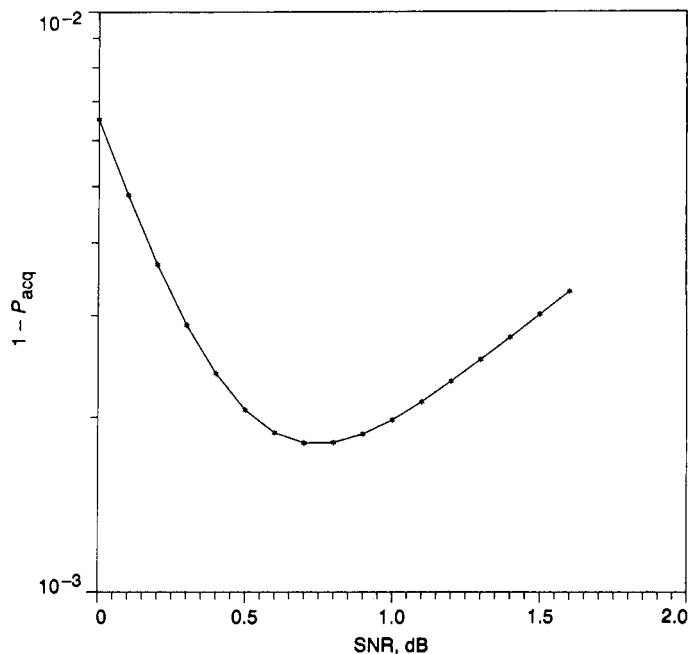


Fig. 7. (15,1/4) code-acquisition performance of scheme 3 using a fixed threshold from 0–1.5 dB.

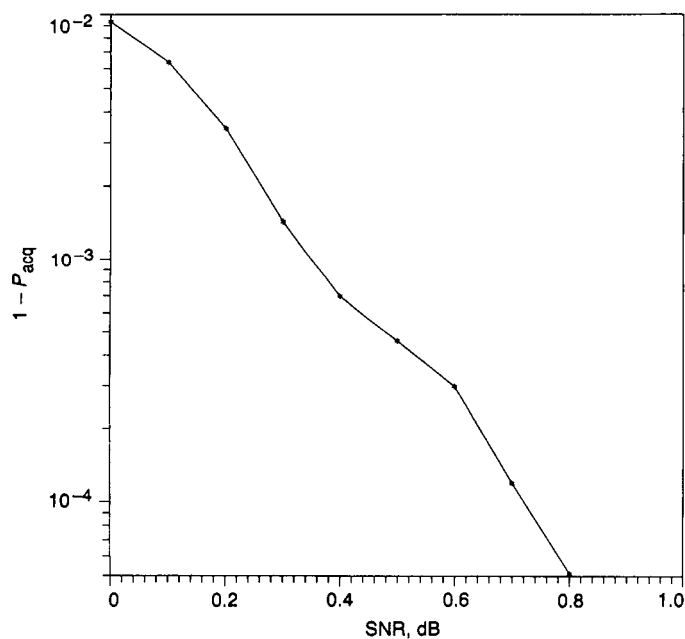


Fig. 6. (15,1/4) code-acquisition performance of scheme 2 using a fixed threshold from 0–0.8 dB.

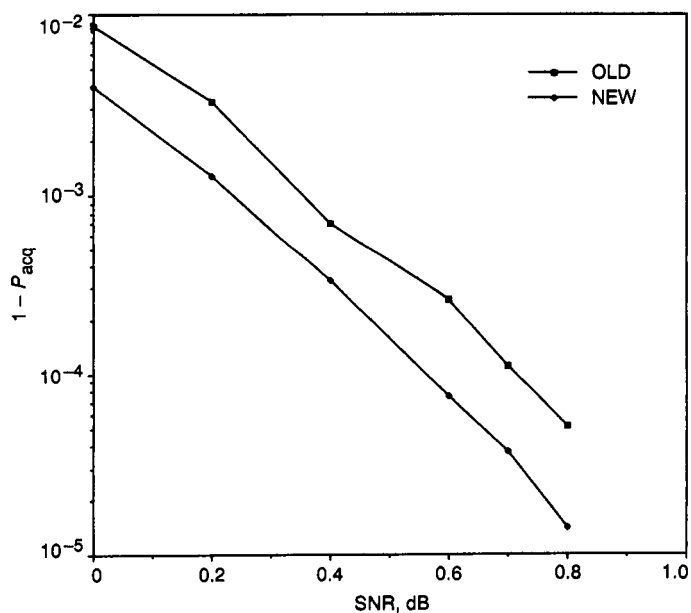


Fig. 8. P_{acq} of the old and new strategies.

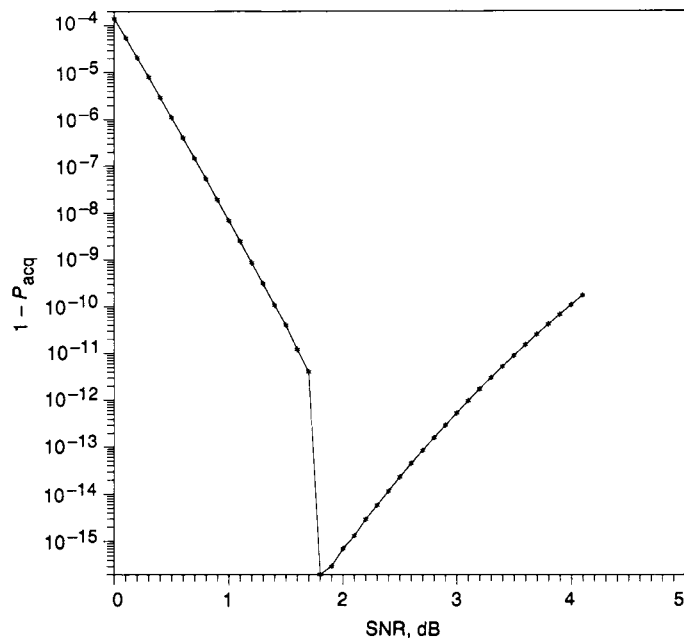


Fig. 9. (15,1/4) code-acquisition performance of scheme 3 using a fixed threshold from 0–4 dB (64-bit sync marker).

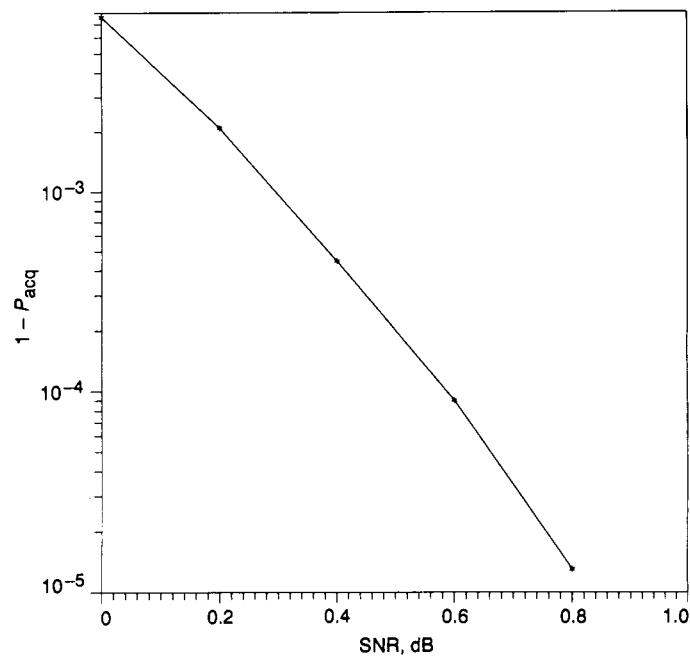


Fig. 10. (15,1/4) code-acquisition performance of scheme 2 using a fixed threshold from 0–0.8 dB (64-bit sync marker).

25-32,
1175
P-13
N92-24316
JJ57 4450
VG 872712

250-kW CW Klystron Amplifier for Planetary Radar

R. Cormier

Radio Frequency and Microwave Subsystems Section

A. Mizuhara

Microwave Power Tube Products, Varian Associates, Inc.

Palo Alto, California

This article describes the design, construction, and performance testing of two Varian klystrons, model VKX-7864A, which replaced the aging and less efficient VA-949J klystrons in the X-band planetary radar transmitter on the Goldstone, California, 70-meter antenna. The project was carried out jointly by the Jet Propulsion Laboratory and Varian Associates. Output power was increased from 200 kW to 250 kW continuous wave per klystron, and full dc beam power is dissipated in the collector (it was not possible to operate the VA-949J klystrons without RF drive because of limited collector dissipation capability). Replacements were made with a minimum of transmitter modifications. The planetary radar transmitter is now operating successfully with these two klystrons.

I. Introduction

The X-band planetary radar transmitter at the Deep Space Network tracking facility at Goldstone, California, has been in operation since 1974. The planetary radar provided the first topological map of planet Venus. Valuable information has also been gained from radar echoes obtained from the planets Mercury, Mars, Jupiter, and Saturn, from the moons of Jupiter and Saturn, including the rings of Saturn and possible rings of Earth (see this issue), and from nearby asteroids.

The transmitter contains two klystrons operating in parallel to generate the output power. In 1986, a study contract was placed with Varian for a klystron of new design, Varian's model VKX-7864A, to replace the aging, less

efficient, low-reliability, and lower power klystrons, Varian's model VA-949J. In April 1991, installation of the two newly designed klystrons upgraded the transmitter system to an output power of 450 kW and an operating center frequency of 8510 MHz.

The VA-949J has a yoke-type electromagnet. Since the collector coil slips over the collector during mounting, the collector diameter is limited. Dissipation of full dc beam power requires a much larger collector, which requires redesign of the electromagnet to a solenoid. With a solenoid-type electromagnet, the klystron is installed gun-first, which allows the collector to be independently sized. The disadvantage is the extreme dimensional accuracy required in construction of the solenoid to realize alignment

between the magnetic field and the electron beam. As a result, the klystron body must be as short as possible, but not so short that the required gain and efficiency cannot be met.

The increased klystron output power to 250 kW places severe demands on the cooling of cavities and cavity tuners, especially output and penultimate cavities. The initial paper design had an extended interaction output circuit, but at the end of the design phase, Varian concluded with JPL concurrence that for this application a single-gap output circuit using an innovative cooling approach was simpler and more cost-effective. The addition of a prepenultimate cavity and the tuning of cavities for maximum efficiency resulted in a 45-percent efficiency.

II. Description of the VKX-7864A Klystron Amplifier

Design requirements for this klystron amplifier tube are a saturated gain of 50 dB, a bandwidth of 40 MHz (-1 dB points), and an electronic efficiency of 45 percent. The tube delivers 250 kW at a frequency of 8510 MHz, which is 15 MHz higher than the previous operating frequency of the planetary radar. Operating frequency band edges are from 8500 to 8520 MHz, with a power variation over the 20 MHz of 0.25 dB. A beam-current requirement of 11.5 A maximum was constrained by existing equipment at the tracking facility. The tube was required to fit into the existing socket and cabinet assemblies used for the VA-949J.

Primary emphasis has been on improving reliability of the klystron amplifier. A reduction of maximum electric field gradient in the gun region from 140 kV/cm to 90 kV/cm eliminated arcing problems experienced in the VA-949J. Decrease of cathode loading and coating the cathode with a binary alloy lowered cathode temperature and the barium evaporation rate by a factor of four. Tube length was minimized to reduce size and weight, since a longer tube requires a heavier focusing solenoid assembly. The VKX-7864A has a Pierce-type electron gun, six cavities, and a collector. Since the available beam supply is current limited, the klystron gun was designed to operate within the limits of this supply.

Figure 1 is a photograph of the klystron mounted in its solenoid and sitting in its stand. The klystron has an appendage vacuum pump for monitoring gas pressure inside the tube.

A. Electron Gun

Figure 2 is a computer-simulated model of electrostatic electron trajectories, and shows a final cathode-to-beam area convergence of about 37:1. The figure is a sectional view of the axially symmetric electron gun with only the upper half shown. Electrons emitted from the cathode converge at a place identified on the figure as "minimum beam radius." Equipotential lines are shown transverse to electron paths. Figure 3 is similar to Fig. 2 except that electric field gradients were taken as shown at points A, B, and C. From the electric field distribution around the axially symmetric diode gun, a maximum gradient of 90 kV/cm occurs on the focus electrode at point A in the figure. The earlier VA-949J klystron had a maximum gradient of 140 kV/cm and consequently required very long dc aging periods.

Having produced an electrostatic beam of correct purveyance and minimum diameter, the next step was the design of a magnetic field to maintain the beam with an acceptable degree of scalloping, and verification testing in the beam analyzer. A scaled model of this gun was fabricated for testing in the beam analyzer. The beam analyzer gives a cross-sectional measurement of current density along the axis of the actual electron beam. Minor adjustments of magnetic field reduced scalloping from 12 percent to 7 percent. Figures 4 and 5 show beam-analyzer measurements of the current density profile at various axial distances from the gun pole piece aperture reference plane (in the Z-axis direction).

Figure 6 shows two views 180 deg apart of a computer-drawn perspective of the beam's cross-sectional current density based on beam-analyzer data. It exhibits normal current density tapering from 4 A/cm² at cathode edge to 2 A/cm² on axis. Minor perturbations in current density indicate patchy emission quality of the cathode's emitting surface near its axis. These perturbations disappeared with proper aging.

B. Klystron Body

The new klystron consists of an input cavity, four buncher cavities, and an output cavity. The RF signal to be amplified at the input cavity starts the velocity modulation process. The field free space between cavities allows the beam to drift and form bunches. Each subsequent cavity couples to the beam, reinforcing bunching energy. The output cavity couples energy from the beam to the output waveguide. Cavities 1, 2, and 3 are stagger-tuned to achieve the required flat bandpass response. Cavities 4 and 5 are tuned for efficiency at a frequency above the high-band edge, enhancing bunching action of the beam.

Cavity 6, the output cavity, is always tuned to center frequency. A schematic view of the six cavities is shown in Fig. 7, and typical buncher cavity dimensions are shown in Fig. 8.

A coaxial-to-waveguide transition with a matching coupling iris between waveguide and input cavity provides a low voltage standing-wave ratio (VSWR) load to the input signal. A ceramic vacuum window with a coupling probe isolates the vacuum of the input cavity from the air side of the coaxial coupler. The probe converts the TEM coaxial mode into the TE₁₀ waveguide mode for coupling to the input cavity. The size of the coupling iris matches the impedance of the cavity to the waveguide. For a cavity gap spacing of 1.0 radian, a beam Q of 321, and a cold unloaded Q of 3500, the external Q equals 294. Use of a coaxial input makes it easier for the input line to exit through the solenoid-type electromagnet.

The beam coupling coefficient for a gridded gap $M = \sin(\omega T_o/2)/\omega T_o/2$ measures the ratio of ac current induced in the cavities to the ac component of the electron beam. For most klystrons, the gaps are not gridded and the beam coupling coefficient is more complex, involving electric and magnetic field geometry at the cavity gap. The ratio of gap spacing to electron velocity, $d/v = T_o$, determines the time electrons remain in transit in the cavity gap. The product of the characteristic impedance of cavities (R_{shunt}/Q) and the beam coupling coefficient squared was maximized to yield the specified saturated gain of 50 dB; both are dependent on cavity gap spacing. Characteristic impedance of a cavity increases as gap spacing increases, while the beam coupling coefficient decreases with increased gap spacing. For this klystron, the value of the coupling coefficient equals about 0.7 and the characteristic cavity impedance equals 108 ohms. The optimum product of characteristic impedance and the beam coupling coefficient squared gives a gap spacing of 1.2 radians.

The final design is a compromise to meet or exceed all design goals simultaneously: RF power output, gain, bandwidth, and efficiency. From proprietary small-signal and large-signal computer programs, spacing between cavities as well as cavity external Q and frequency were found. Tuning based on calculated small-signal gain response, shown in Fig. 9, was later modified in hot test to optimize output power and efficiency.

The efficiency depends mostly on the tuned frequency of prepenultimate and penultimate cavities. A prepenultimate cavity complicates the problem of finding optimum tuning. The process involved systematic mapping of large-signal RF response by holding the frequency of one cavity

constant while varying another. This also must be done with cavity 3 as a parameter. The efficiency of the fabricated tube measured 45 percent, the design value specified.

Temperature variations of the buncher cavities cause the frequency of cavities to change. A cold test of a cavity measured a straight line slope of -0.158 MHz/deg C. Because these tubes at the transmitter site in the Mohave Desert experience a normal coolant temperature variation of about 30 deg C from midmorning to early afternoon, cavity frequency can change at a rate of 0.5 to 1 MHz per hour. Corresponding phase change would be intolerable, especially when the transmit-receive cycle may be several hours. Consequently, it is necessary to have a dedicated, temperature-regulated coolant supply for the klystron RF (cavity) circuit and output waveguide.

Thermal design is a critical part of klystron design. Very tight coupling of cooling water to cavities is necessary to remove heat generated by the RF on the walls of cavities. Although cavities are not tunable in operation, during initial testing at Varian the cavities required tuners. Use of a diaphragm tuner on the cavity wall complicates the cooling design. In fact, the initial prototype tube failed because of inadequate cooling of the penultimate cavity diaphragm tuner. The solution consisted of an increase in thickness of the diaphragm to provide better heat conductivity and a modification of the cooling pipe to provide better thermal coupling to the diaphragm.

As a result of bunching, minimum electron velocity is an important parameter in preventing excessive body intercept current. Typically, it is prudent to design for the slowest electron velocity to be no lower than 20 percent of dc beam velocity. During testing, body current (beam current intercepted by the body circuit, which is primarily at the output cavity) tripled when phase modulated. Originally, the body current overload was to be set to 30 mA but, because of the increase in current due to phase modulation, a new value of 60 mA was selected. However, when measured calorimetrically, both with and without modulation, the body power remained nearly unchanged, which implies that the energy of additional electrons intercepted with modulation is negligible. This increase in body current is believed to result from those sidebands (created by the modulating process), which add in-phase to generate a higher gap voltage. A higher retarding gap voltage would tend to strip off slower electrons in the bunched beam at the output cavity. The drawback of this increased body current is the necessity of setting body current interlock protection to twice the preferred value. This increases the risk of potential damage from overheating due to a fault when the carrier is not modulated.

C. Window

The output window is a thick-block (typically a half-wavelength) round window. The material is beryllium oxide ceramic. A cooling jacket surrounds and supports the window. Output VSWR return loss versus frequency, shown in Fig. 10, is the result of measured data on a window before the window was mounted to the klystron. With 250 kW of output power, window dissipation is 800 W.

D. Collector

The collector is an oxygen-free high-conductivity copper cylinder with a closed top. The inside diameter is 20 cm. Liquid coolant passes through holes in the walls to cool the 2.5-cm-thick sidewalls. The beam spreads in the collector due to the repulsive force of electrons on each other. Lack of a magnetic field allows electrons to spread freely. When there is no RF on the beam, full beam power must be dissipated in the collector. From Fig. 11 the peak power density for the dc beam is 670 W/cm. This occurs where the beam first impinges on the collector. The intercepted beam power has a narrow and very sharp peak, which allows the copper to conduct heat to surrounding areas. Nevertheless, wall temperature at the cooling passage could cause coolant boiling if coolant pressure were near atmospheric. To prevent boiling, this area is connected to the inlet or high-pressure side of the heat exchanger (nominally 8.5 kg/cm²). When the klystron is producing RF, collector power is reduced by the RF output power and the beam spreads faster than a dc beam. Although some electrons have greater velocities upon entering the collector, the majority have lower velocities. Slower electrons tend to impinge the collector surface in a zone preceding the point where a dc beam first strikes. With an output power of 270 kW, the maximum peak power density is 120 W/cm².

The collector end cooling device contains spiral grooves. Figure 12 shows a cross section of the collector assembly. In the final assembly, a thin lead wrap on the collector serves as an X-ray shield.

E. Magnetic Focusing

A reentrant solenoid focuses the electron beam. This type of solenoid is necessary because of the large size of the collector. During installation, the klystron is lowered into the solenoid gun-first, until the collector pole piece rests on the solenoid outer cylinder. The collector is above the electromagnet, and the collector size is completely independent of the solenoid. The solenoid requires tight dimensional accuracy and serves to shield against any X-rays generated in the klystron body.

The minimum field required to focus a dc beam of constant axial velocity, current density, and radius is called the Brillouin field. Because of the need to keep intercepted body current to an absolute minimum under RF conditions, the magnetic field must be about 2.7 times the Brillouin value. The magnetic field leaks into the cathode region for convergent flow focusing. Except for the small leakage field in the cathode region, construction of the solenoid is such that it is fully magnetically shielded. The solenoid is 46 cm in diameter and 23 cm high and weighs about 122 kg. The total weight of klystron and magnet is 333 kg.

III. Performance Measurements

Test data measured on the first two klystrons fabricated have characteristics so close to each other that with both tubes operating from a common beam power supply, transmission and combining losses are within 0.5 dB. The data presented here are not from one klystron. Some test results are from klystron Serial Number 101, while other results are from Serial Number 102.

A. Bandpass

The klystron operating frequency is from 8500 to 8520 MHz, and bandpass data, shown in Fig. 13, are limited to this frequency range. Bandpass curves are for beam voltages of 40 to 51 kV, with the drive adjusted to provide saturated output at each value of beam voltage. At 51 kV the output power at 8510 MHz is 260 kW with a variation of -0.25 dB at band edges. Because the output cavity is for all practical purposes untunable, the resonant frequency of the output cavity turned out to be about 7 MHz lower than the desired center frequency of 8510 MHz. However, required minimum output power of 250 kW at 8510 MHz was still met with a 51-kV beam because of the relatively flat response.

As previously stated, one problem experienced was an increase of body current with phase modulation. With a modulated signal at the input to the klystron (Fig. 14), the switching of modulation from on to off causes an abrupt change of body current, as seen in Fig. 15. With modulation, body current was 42 mA at a carrier frequency of 8500 MHz; at 8510 MHz when modulation was turned off, body current suddenly decreased by 20 mA. Figure 15 shows that there was no change of output RF power at the time of switching modulation. The switching of modulation and the accompanying sudden change in body current was observed while slowly sweeping the frequency from 8500 to 8520 MHz. Subsequently, the klystron body

was instrumented for measuring power dissipation and the power does not change with or without modulation. This result confirms that increased body current does not contribute additional power dissipation in the cavities of the klystron.

B. Group Delay and Phase Linearity

Group delay remains practically unchanged with the klystron operating saturated or unsaturated, as shown in Fig. 16. Group delay data taken at a beam voltage of 53.5 kV, which is 2.3 kV above normal operating value, are also shown in Fig. 16. This value of voltage was used to examine the sensitivity of group delay to variations of beam voltage. The group delay remains practically unchanged with variations in output power and operating beam voltage. Variations are attributed to instrument noise. The output phase linearity is shown in Fig. 17.

C. Phase Pushing Factors

Phase pushing factors are listed in Table 1. The coolant temperature caused a larger change in phase than what was specified. During initial testing at Varian, a common coolant line for the klystron body and waveguide was used, and phase changes were associated with temperature changes of klystron body and waveguides. Varying the collector coolant temperature had no measurable effects on phase. For the WR-125 waveguide operating at 8510 MHz, the phase changes with a temperature equal to 0.21 deg/deg C per meter of waveguide length. It is known that the phase pushing factor is negative for the klystron and positive for the waveguide. The coolant is stabilized to better than ± 1 deg C; consequently, the phase pushing factor, although much greater than originally specified, is not a problem.

IV. Conclusion

The first two VKX-7864A klystrons tested met the requirements as replacements for the VA-949J klystrons in the planetary radar transmitter. These new tubes are now operating in parallel at Deep Space Station 14, Goldstone, California. Combined output power from the two VKX-7864A's is about 500 kW, but because of losses due to combining and waveguide transmission, radiated power at the feedhorn is 450 kW. This value of radiated power is about 100 kW more than that available from the VA-949J's. The collector of each klystron is able to dissipate full dc beam power, which greatly simplifies radar detection and analysis of near-Earth asteroids and man-made debris.

These two tubes have very close characteristics, allowing operation from a common beam supply. Both tubes have nearly identical bandpass and gain responses as well as group delay and phase linearity across the band of interest. The phase pushing factors as a function of various operating parameters met their requirements in all cases except for that of body coolant temperature. Use of a temperature-regulated coolant source for body and waveguide would have been required even with the lower pushing factor. The higher pushing factor of the body is acceptable.

The only other unexpected phenomenon, increased body current with phase modulation, proved to be insignificant because the intercepted electrons have little kinetic energy and do not add to body thermal dissipation. These klystrons represent state of the art for X-band CW klystron amplifiers.

Table 1. Klystron phase pushing factors.

Parameter	Limits	JPL measurements	Varian measurements
Beam voltage, deg/V	0.04	0.023	0.022
Drive power, deg/dB	5.0	2.67	1.6
Magnet current, deg/percent	2.0	— ^a	0.7
Coolant temperature, ^b deg/deg C	0.10	0.87	1.2
Filament current, deg/percent	0.4	— ^a	— ^a

^a No measurable relationship.

^b Varian measurement varied body and external waveguide temperature together; JPL measurement varied body and external waveguide temperatures independently (note large deviation from limit).

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

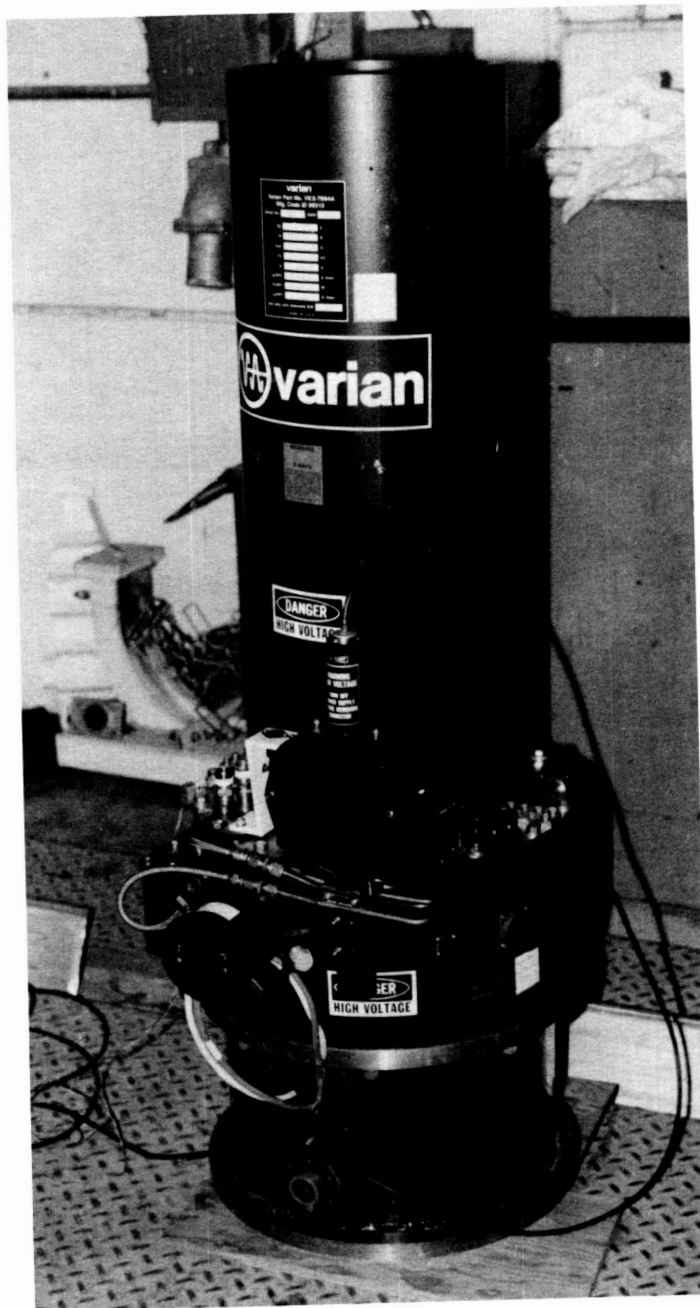


Fig. 1. Klystron and solenoid assembly, Model VKX-7864A.

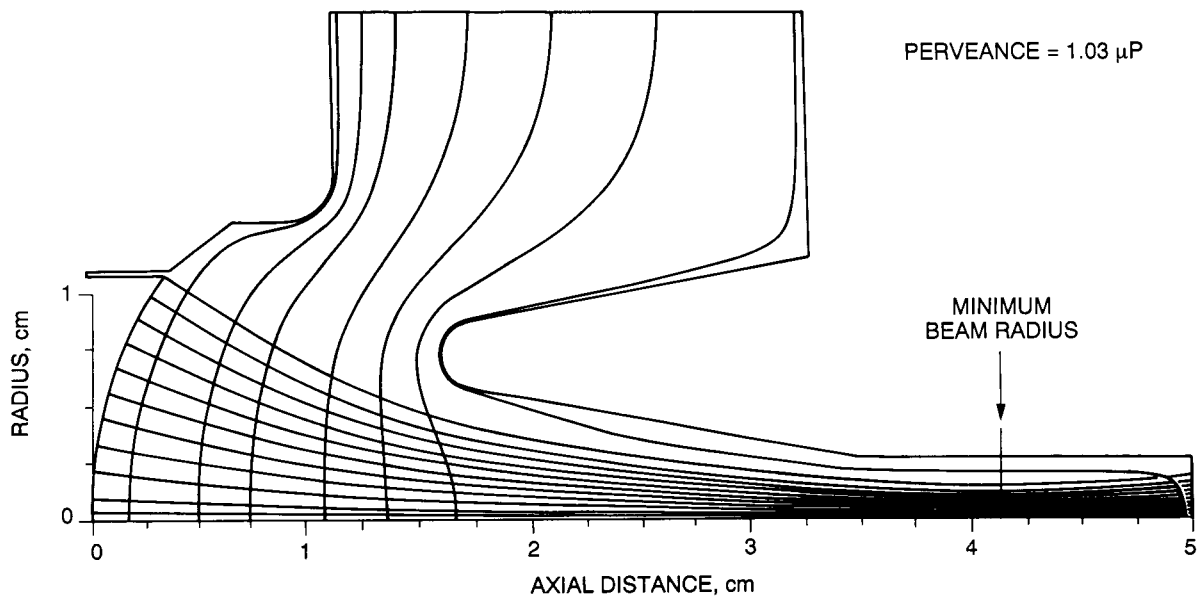


Fig. 2. Typical computer-simulated electrostatic electron trajectories.

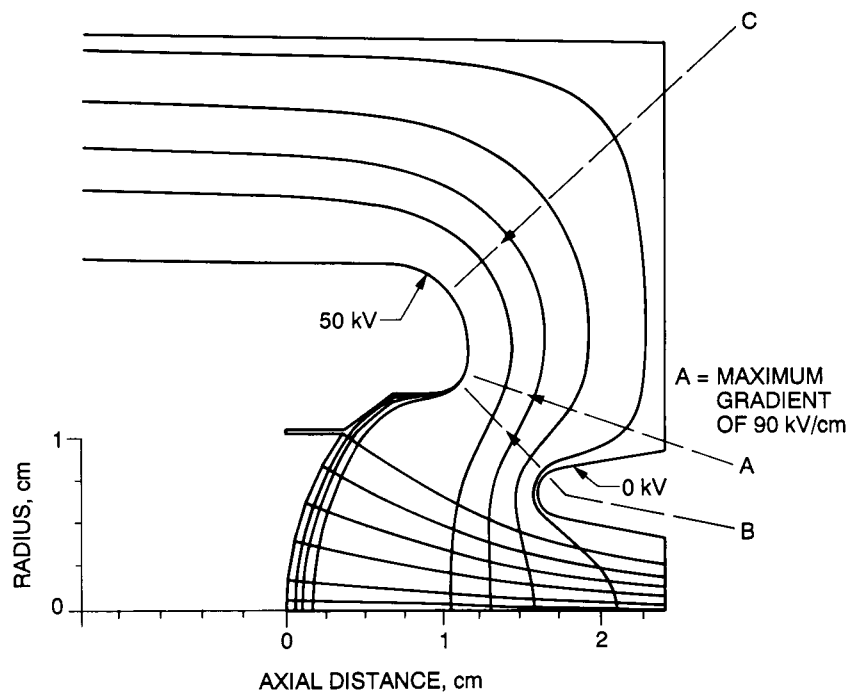


Fig. 3. Equipotential lines between cathode and anode and highest gradient point.

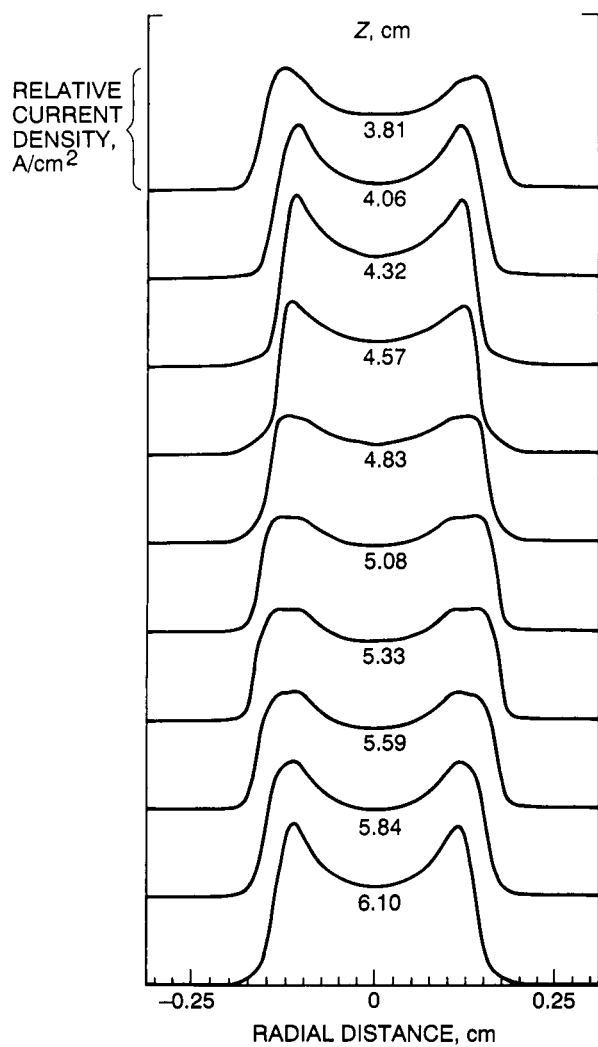


Fig. 4. Beam analyzer measured current density profile at various axial positions from 3.8 to 6.1 cm (confined-flow magnetic field).

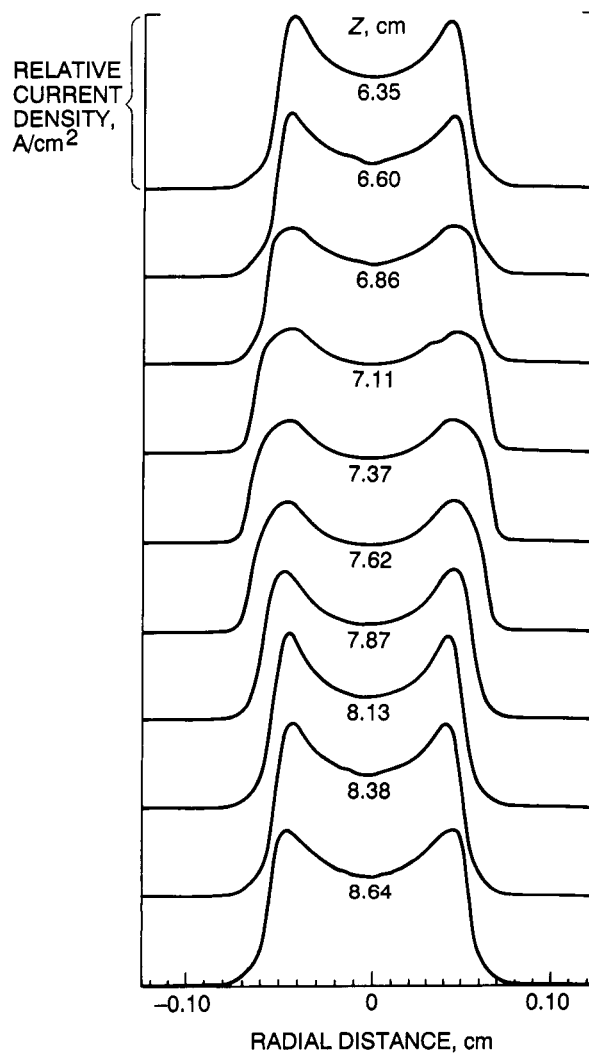


Fig. 5. Beam analyzer measured current density profile at various axial positions from 6.4 to 8.6 cm (confined-flow magnetic field).

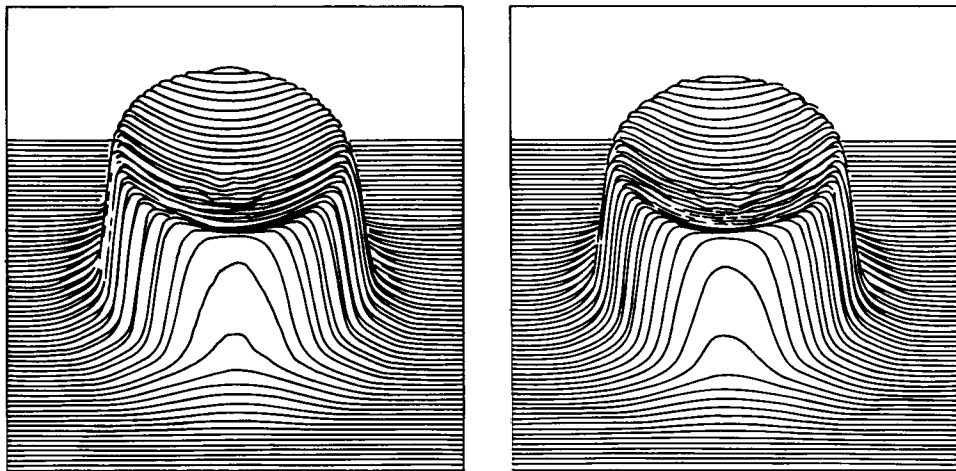


Fig. 6. Three-dimensional beam density profile viewed from two positions 180 deg apart.

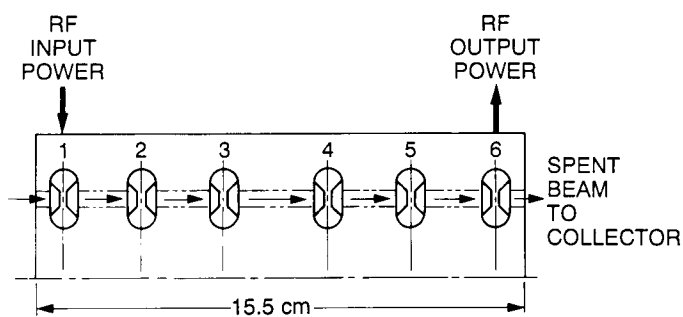
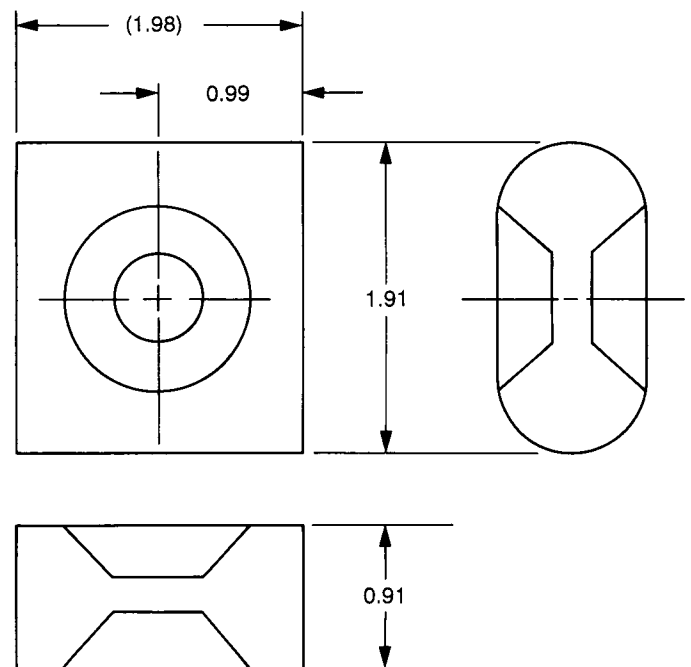


Fig. 7. Schematic view showing cavity location.



ALL DIMENSIONS IN CENTIMETERS

Fig. 8. Typical buncher cavity size.

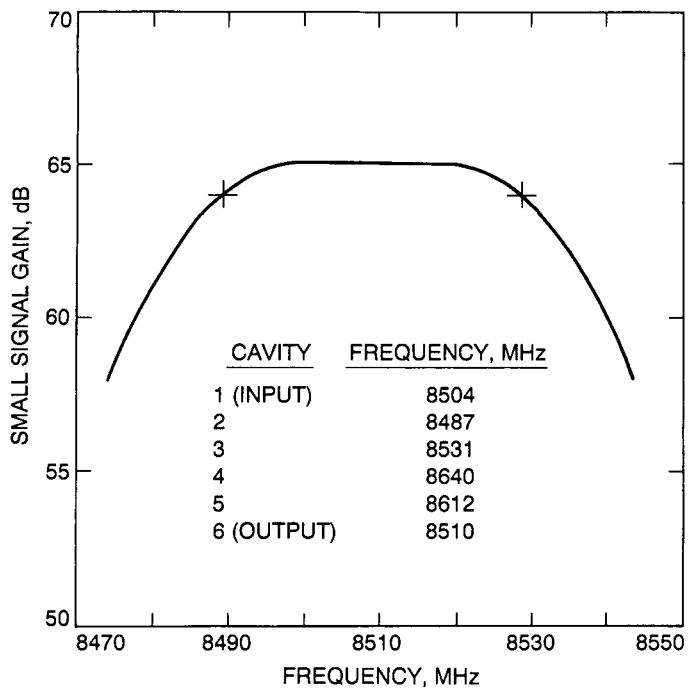


Fig. 9. Small-signal gain response (calculated).

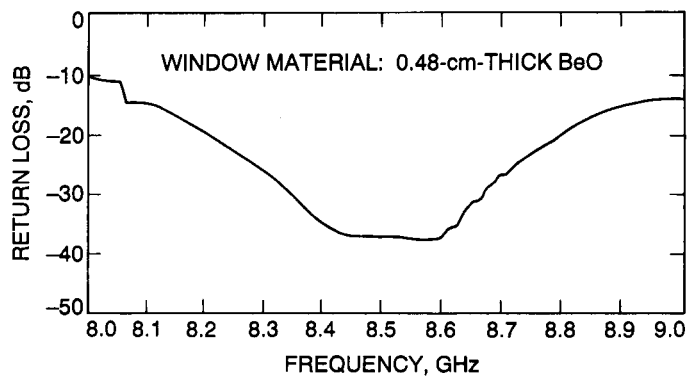


Fig. 10. Output window return loss versus frequency.

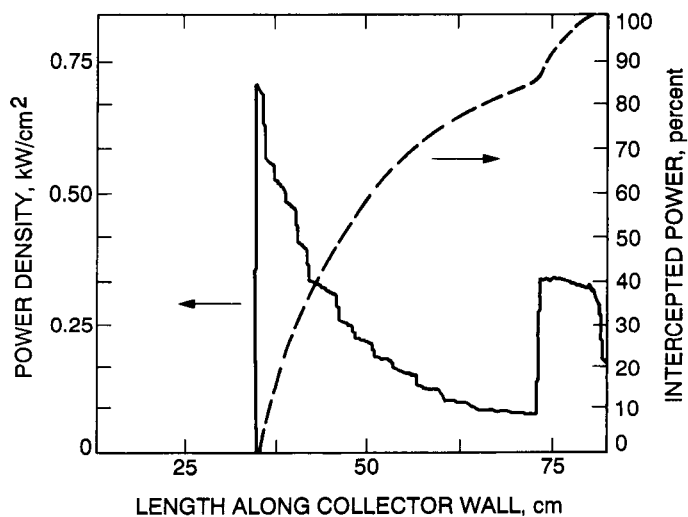


Fig. 11. Collector power density—dc beam.

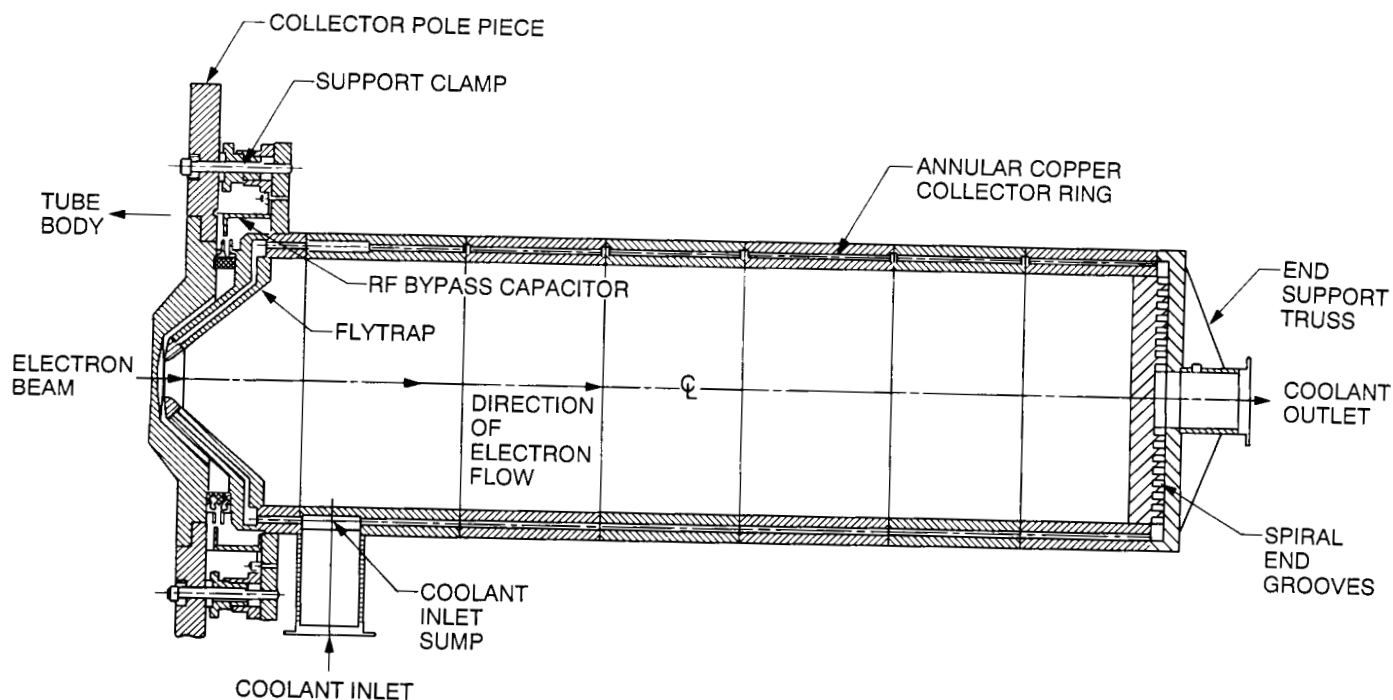


Fig. 12. Collector layout—cross section.

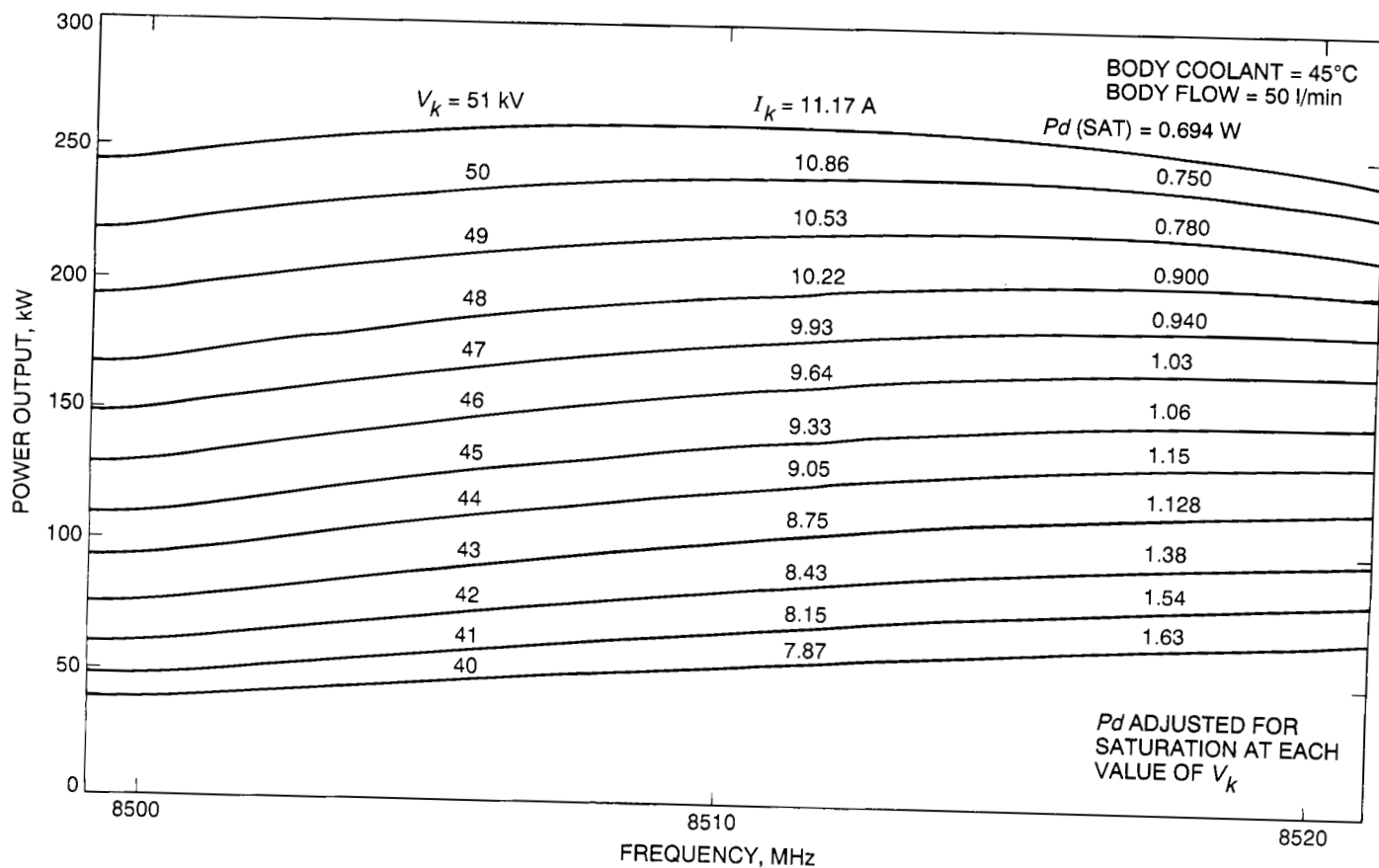


Fig. 13. Power output versus frequency.

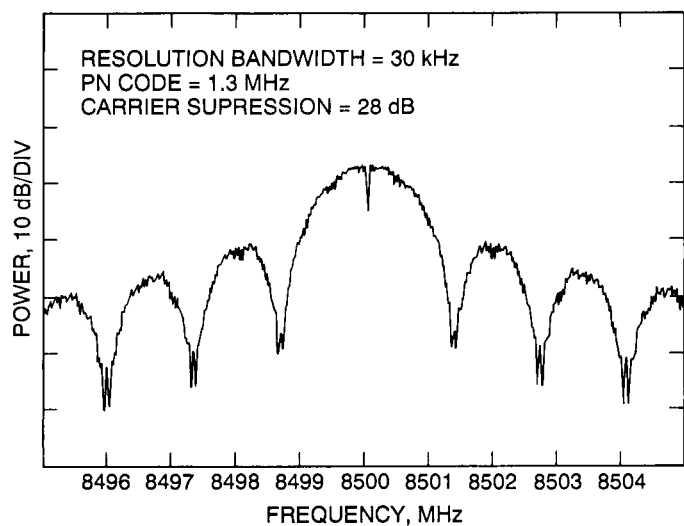


Fig. 14. Phase modulated carrier.

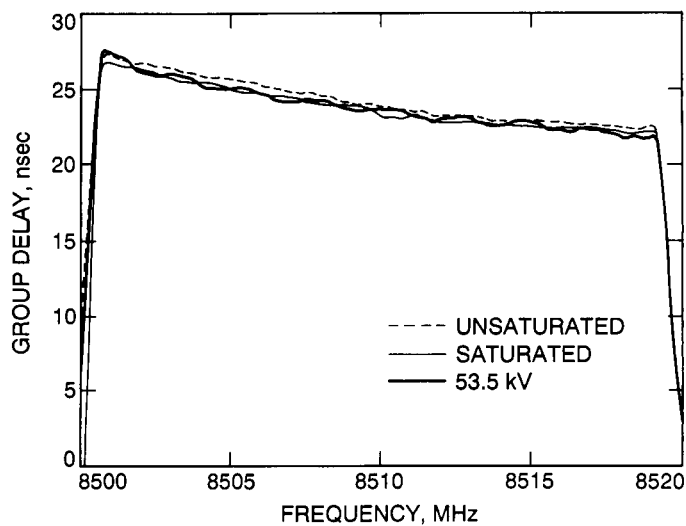


Fig. 16. Group delay.

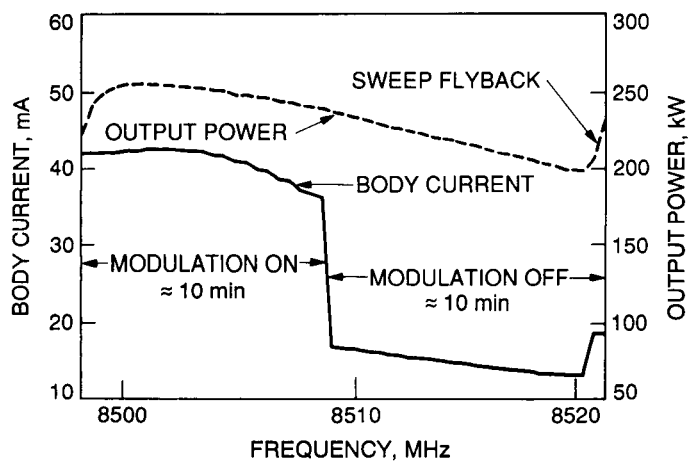


Fig. 15. Body current with and without modulation.

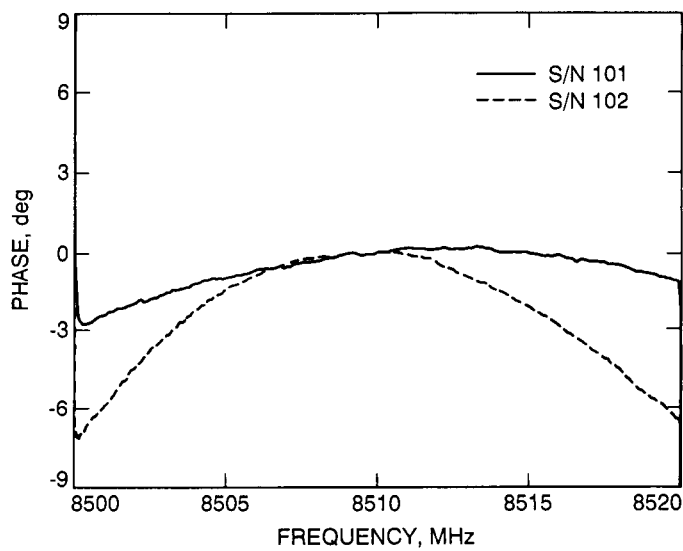


Fig. 17. Phase linearity.

510-2-
70076 N92-24317
p-8

JJ 5-20-92

An Assembly to Provide a Time Reference for the JPL Network Operations Control Center, Real-Time

P. A. Clements and J. G. Lopez
Communications Systems Research Section

M. J. Kelley
Communications Systems Research Section and Telos Corporation, Pasadena, California

This article describes a device which provides the value of date-coordinated universal time (date UTC)¹ to the JPL Network Operations Control Center, Real-Time (NOCC-RT) facility. The NOCC-RT is the real-time portion of the NOCC upgrade task. The time scale is generated in the NOCC-RT clock processor; however, there is a continuous reference to UTC, as realized by the National Institute of Standards and Technology and transmitted by Earth-orbiting satellites. An important functional design requirement is the 99.9-percent availability.

I. Introduction

The NOCC-RT clock processor (NCP) is a device which supplies the value of date-coordinated universal time (date UTC) to the JPL Network Operations Control Center, Real-Time (NOCC-RT) facility. The NCP is a separate device from the other hardware in NOCC-RT. The software in the NCP is a separate program set, designated NOCC-RT clock reader (NCR), which is located in programmable PROM's on a single-board computer. The entire device is designated a DSN assembly of the NOCC-RT Frequency and Timing (NFT) Subsystem. The NFT is one of the seven subsystems which compose the NOCC-RT facility.

NOCC-RT requires that the value of date UTC be available with a precision of 1 msec and an accuracy of 10 msec within the program sets which compose NOCC-RT. The availability of the value of date UTC at the output of the NCP is required to be 99.9 percent. This time scale must be generated from a 1-MHz signal originating at a cesium-beam frequency standard located at the JPL Standards Laboratory (JSL). To ensure that the output is to the required accuracy and precision, it is referenced to the UTC generated at the National Institute of Standards and Technology UTC (NIST) in Boulder, Colorado.

Access to UTC (NIST) is via time transmissions from two Geostationary Operational Environmental Satellites (GOES), which are operated by the U.S. Department of Commerce. UTC can be realized with a precision and accuracy of 100- μ sec using a GOES timing receiver. The

¹ "The reading of a specified time scale. Note that the date can be conventionally expressed in years, months, days, hours, minutes, seconds, and fractions thereof . . . "[1]

accuracy of the NCP was designed to be 1 msec. This allows the 10-msec accuracy requirement of NOCC-RT to be met even with the reduction of precision and accuracy in the computer system. Tests have shown the value of date UTC within NOCC-RT to be within 5 msec of UTC (NIST).

Availability was the most difficult design requirement to meet. The design approach taken was to use redundancy as well as a diversity of sources for the value of date UTC: Both the East and West GOES are used, in concert with a local cesium clock. The value of date UTC is received from the three sources, then compared. If two or three agree, that date is accepted; if all three disagree then the date of the local clock is the date accepted, but fault flags are set in the status bytes of the output to indicate uncertainty. These flags correspond to the lights on the front panel of the computer.

II. Design Process

A. NOCC-RT Frequency and Timing Subsystem

The design of the NFT was derived from functional requirements.² These requirements were combined with requirements for the remaining six NOCC subsystems which compose NOCC-RT to develop the functional design document (FDD) for NOCC-RT. The FDD was written as a task document and, as such, contained the design for all of NOCC-RT in that particular design phase.

The NCP is largely a hardware device but with significant software. Because the NCR is a program set in a piece of hardware separate from the rest of NOCC-RT, a separate ensemble of design documents was developed starting with the software requirements document (SRD).

B. NOCC-RT Clock Processor

After being separated from the rest of NOCC-RT and at the start of the software requirements analysis phase, the NCP design was divided into hardware and software. The hardware design followed the traditional DSN design practice, which culminates in a hardware transfer agreement. At the same time, documents³ were developed for the software, starting with the SRD and ending with the software transfer agreement.

The hardware and software aspects of the design were remerged primarily in the software test plan and procedures. The NCP was tested as a DSN assembly and the results recorded in the software test report. An interesting example of this remerging was the section of the software test results report which proved the availability requirement had been met. The availability analysis dealt only with the hardware part of the NCP. However, the resulting report is a compact and complete verification that the NCP meets the hardware and software functional requirements assigned to it.

After the NCP was tested in place, it was connected to the NOCC-RT computers. The entire NCP was then included in the NOCC-RT integration testing, which was at the NOCC upgrade task level.

III. Hardware Configuration

The NCP consists of two GOES receivers, a time-code translator/generator (TCT), a computer with a front panel containing status lights and a fiber-optic interface for the antenna cables. Figure 1 shows the NCP installed in the NOCC-RT cabinet at the JPL Space Flight Operations Facility (SFOF). Figure 2 shows a schematic representation of the NCP. Notice that the NCP receives a parallel time code from the three sources and sends a serial time code to NOCC-RT.

The two GOES receivers are configured using front-panel switches. The East and the West GOES signals are selected by a toggle switch. The location of the receiver is entered using a thumb-wheel switch. The satellite's signals are received by antennas located on the roof of a building which is adjacent to SFOF. The output of the receivers is a parallel time code with millisecond resolution. A serial status message is also sent to the NCP. The status message is used to build the status bytes, which are included in the serial time code output from the NCP.

The TCT can properly function in two different ways in the NCP: It can translate a serial time code to a parallel time code or it can generate a parallel time code from a source of 1 MHz. It is configured to accept a 1-MHz signal which is originated at the JSL. The JSL provides time and frequency service anywhere at JPL usually using a directly hard-wired 600- Ω phone line. In the case of the NCP, the signal is proximately provided by a 1-MHz distribution system internal to SFOF. Again, the output of the TCT is a parallel time code, with a format identical to that of the GOES receivers and the stability of a cesium frequency standard.

² *Network Operations Control Center Subsystems Functional Requirements: Frequency and Timing Monitor Subsystem (1991-1995)*, JPL D-5423 (internal document), Jet Propulsion Laboratory, Pasadena, California, January 15, 1989.

³ *JPL Software Management Standards Package*, JPL D-4000 (internal document), Jet Propulsion Laboratory, Pasadena, California, December 1988.

The computer is a Multibus I configuration with a CPU board, three I/O boards for the parallel time input and status input from the GOES timing receivers, and a special board to operate the front-panel status lights. A commercially available chassis is used, with new front and back panels added. Figure 3 shows the computer with the front panel opened up. The CPU board is on top. The next three boards are I/O to the time-code sources and the bottom board operates the front panel.

The front-panel status lights provide the same information that is sent in the status bytes of the output message. Green lights indicate correct functioning while the red indicate fault. The indicators are divided into two groups. The first group shows the condition of the inputs to the NCP. For example, if the value of date from the satellite receiver is unreliable because of a loss of signal, a red light will switch on. The second group indicates the condition of the output of the NCP. The NCP will accept the date UTC, but if it judges the date to be inaccurate, a red light will switch on.

There are instructions on the front panel and inside the front door so that installation or replacement of equipment can be accomplished without the installation manual. The switch positions on the front of the receivers are given on the computer's front panel. The cabling diagram is affixed to the inside of the front door.

Because of SFOF electrical conductivity requirements, part of the GOES antenna cable was replaced with fiber-optic cable. Special conversion modules were designed, constructed, and used in the antenna line. The configuration is shown in Fig. 4. The optical fiber provides a nonconductive path through the boundary of SFOF and still allows normal operation of the GOES receiver.

IV. Software Structure⁴

Being able to read a clock and report the value of date to a computer requires knowing the amount of time that each task takes to complete. Also, the NCR is required to respond frequently and quickly to requests for date UTC. These requirements led to the decision to write the software in assembly language. This was not a difficult problem because the program was small. The use of assembly language also allowed the application program to handle

its own interrupts, which in turn allowed precise control of the amount of time needed to read the clocks.

The millisecond counter measures the time between clock-reading cycles. The accuracy and precision requirements to allow the NCP to read a clock only in response to a request were too complex because the NCP must be ready to send the value of date immediately. Instead, the NCP reads the clock on a regular schedule, and when a request for the value of date UTC comes in, the reading cycle stops, and the NCP sends out the most recent value of date UTC with the accumulated time-interval value from the millisecond counter.

NOCC-RT wanted the time mark to be at the end of the message, and not at the beginning, where it is usually placed. So the time required to transmit the message, plus about 10 milliseconds, is added to the value of date. The accepted value of date is sent; then a pause occurs to monitor the millisecond counter until it reads the value of the date sent. The time mark is then sent on the exact date UTC.

The four major tasks of the NCP are to read the time code from the three clocks, get the status information from the two GOES receivers, count milliseconds, and send the value of date UTC to NOCC-RT upon request.

The two GOES receivers and the time-code translator/generator all produce time code with the same format, a parallel binary code. The three outputs are read in rapid succession, then compared in a "majority-vote" algorithm to obtain a value of date UTC.

Status information on the condition of the date value from the GOES receivers is obtained from serial ports on the receivers. There is no status information from the time-code translator/generator.

The millisecond counters are started, one of them at each reading of the clocks, when the clocks are read. This allows an interpolation of the value of date UTC between clock readings. The clocking pulse interrupts the microprocessor every millisecond to update the counter. The total run time of the millisecond counter is from 20 to 30 msec, which is the clock-reading-cycle time.

Lastly, the NCP services the request for the value of date UTC from the NOCC-RT facility. At the end of each clock-reading cycle, the user-request lines are polled to determine if a request has come in. When a request is recognized, a value-of-date-UTC message is formatted and transmitted to NOCC-RT.

⁴ For a complete description of the software structure, refer to *Software Specification Document (Volume 1), NOCC Frequency and Timing Subsystem, NOCC Clock Reader Software, NOI-5450-OP-A/1.0* (internal document), Jet Propulsion Laboratory, Pasadena, California, 24 January 1992; vol. 2 in press.

V. Interface Between the NCP and NOCC-RT

The interface between the NCP and the NOCC-RT computer is a standard RS-232 format. At the time the original design work was done, it was not known how close, physically, the NCP would be to the NOCC-RT computers. It was decided to provide for a fiber-optic output of the date information. By using a 25-pin-electric to two-fiber-optic-cable interface, the value of date UTC can be transmitted up to two kilometers from the NCP. This means that the value of date UTC can be made available anywhere at JPL with a precision and accuracy of 1 msec. Before the NCP was installed in SFOF, the value of date UTC was provided to operational NOCC-RT in SFOF from the development laboratory, located over 300 m distant, by using fiber-optic cables.

The value of date UTC is requested by NOCC and the NCP responds with the date message. The date message is different than the one usually encountered. The message begins with the value of date (to the millisecond); then there are two bytes of status information. The date-UTC message is valid at the time that the last byte of the message is sent.

Requests for the value of date UTC are from two NOCC-RT computers. Each computer makes a request about once every five seconds. The NCP first sends the value of date UTC and then provides the date mark. The value of date UTC does not necessarily coincide with the second tick, but can be any millisecond of the second.

VI. Installation

The NCP is installed on the third floor of SFOF. The adjacent building (180) has an unobstructed view of the geostationary satellites, so it was decided to install the antennas on that building's roof. The antenna cables were then run down cable ducts from the roof, under the road, into the SFOF basement, and up to the third floor.

While the installation was taking place, SFOF engineering staff decided to not allow the running of cables capable of conducting electricity into the building. A redesign was then started to replace the coaxial antenna cable with fiber-optic cable. The coaxial cable had already been laid from the roof to the second floor of the adjacent building,

so it was decided to leave that cable in place and install a fiber-optic-to-coaxial converter box in the telephone closet on the same floor. This box converts the intermediate frequency and local oscillator signals for the GOES antenna from electrical to optical. The antenna signal was routed through fiber-optic cable to the basement of SFOF and then to the third floor. Another fiber-optic-to-electrical converter box was installed as a part of the NCP hardware to convert the antenna signal back to electrical so that it could be input to the receiver.

VII. Performance

The NCP meets all of the functional requirements assigned to it by the NOCC-RT FDD. The frequency at which the value of date can be received has been tested for up to 17 times per second with an accuracy and precision of 1 msec. Each request is answered with a value of date UTC and a set of validation flags in the status bytes. The value of date UTC includes the proper insertion of leap seconds, as is required in the UTC time scale. Furthermore, notification of approaching leap-second and leap-year events is provided by the flag bits in the status bytes.

All the acceptance tests were done after the NCP was installed in SFOF. The test equipment, including a portable cesium clock to check the accuracy of the value of date UTC, was brought to the NCP to complete the tests.

The NCP will service two computers, as it was designed to do. However, there is no reason it could not service more computers. With the fiber-optic distribution of the value of date UTC upon request, a large system of interconnected computers could stay synchronized by referencing the date UTC at selected points, thus avoiding the problem of reduction of the precision and accuracy of the value of date as it is passed through interconnecting computer structures.

VIII. Conclusions

The NCP is a reliable source for date UTC. The value of date UTC is validated by the NCP and it is a true UTC with all of the leap-second adjustments included.

Reference

- [1] *Recommendations of the CCIR, 1990 (Also Resolutions and Opinions), Vol. VII, Standard Frequencies and Time Signals*, CCIR, Geneva, Switzerland, p. XIV, 1990.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

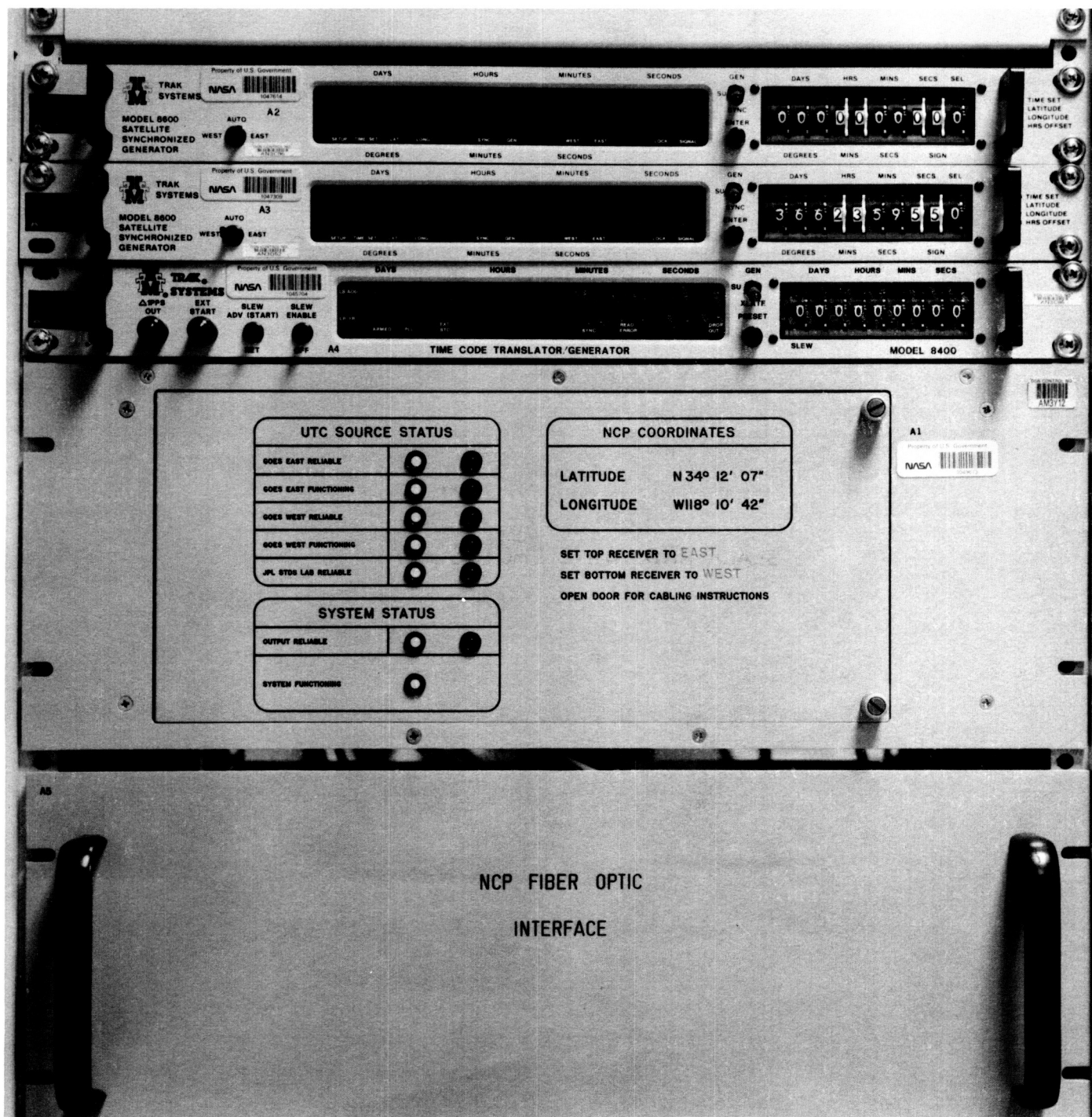


Fig. 1. The NOCC-RT clock processor assembly installed in its cabinet.

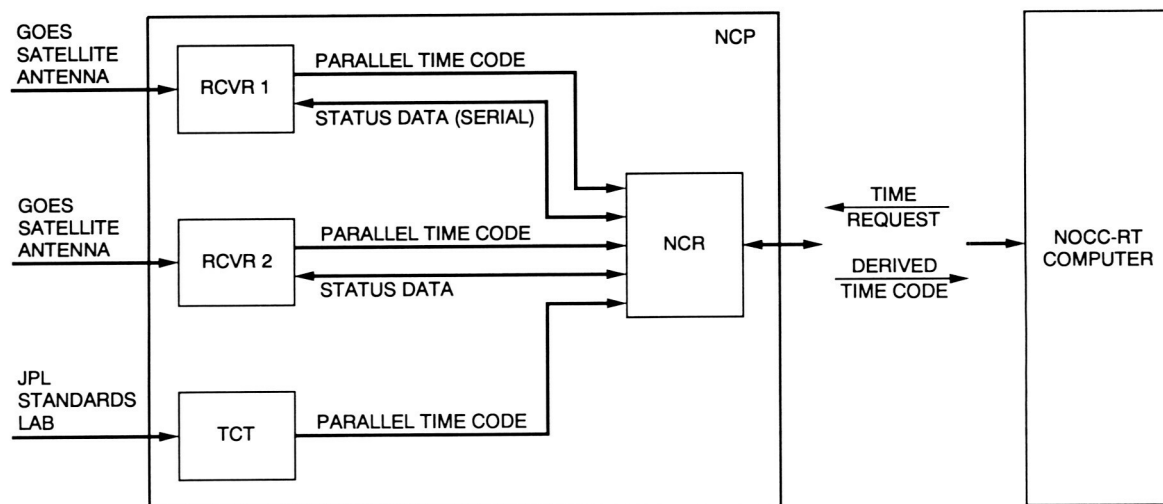


Fig. 2. Schematic diagram of the NOCC-RT clock processor, showing the interconnections between the receivers, the time-code translator/generator, and the NOCC-RT clock reader.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

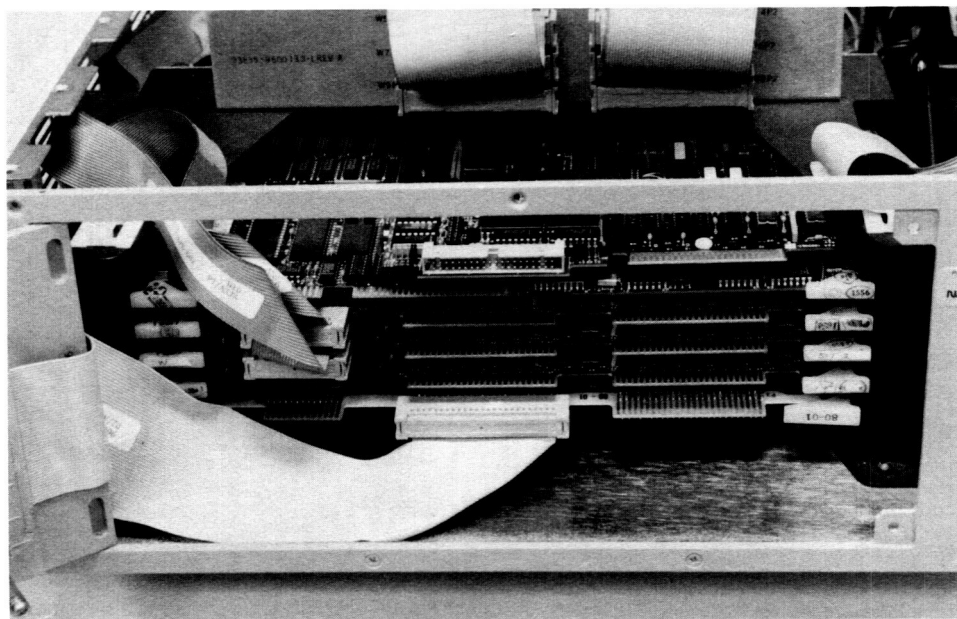


Fig. 3. The NOCC-RT clock reader chassis with the front panel opened, showing the computer and input/output boards.

ORIGINAL PAGE
BLACK AND WHITE PHOTOGRAPH

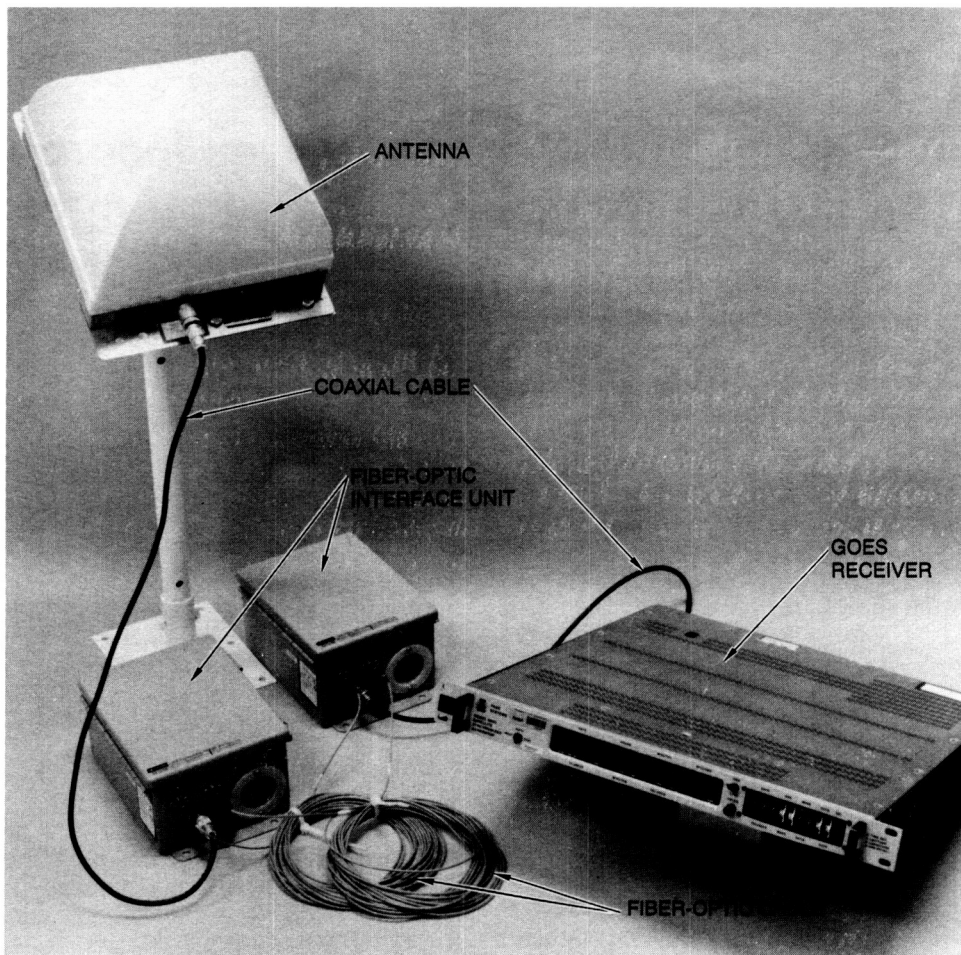


Fig. 4. A GOES receiver assembly.

N92-24318

p-13

JJ 574450

Dynamic Modeling of the Servovalves Incorporated in the Servo Hydraulic System of the 70-Meter DSN Antennas

R. D. Bartos

Ground Antennas and Facilities Engineering Section

As the pointing accuracy and service life requirements of the DSN 70-meter antennas increase, it is necessary to gain a more complete understanding of the servo hydraulic system in order to improve system designs to meet the new requirements. This article develops a mathematical model for the servovalve incorporated into the hydraulic system of the 70-meter antenna and uses experimental data to verify the validity of the model and to identify the model parameters.

I. Introduction

JPL has long maintained the objective of improving the performance and operating life of the DSN 70-meter antennas. A mathematical model of the servovalve incorporated into the 70-meter antenna servo hydraulic system was created to gain a better understanding of the operation of the hydraulic system as one step toward this objective. During the development of the hydraulic servovalve model several approaches were considered. The first approach consisted of physically modeling the dynamic and steady state operation of each of the internal valve components and combining these models as a system to obtain the servovalve model. The internal valve components include the spool, bushing, feedback wire, flappers, nozzles, and torque motor. Due to the complexity of performing this type of analysis and the inevitable uncertainty of the model due to manufacturing tolerances and complex geometries of internal fluid flow paths, this model was determined to be inappropriate. An alternative model that mathematically describes the performance of the servovalve based upon

experimental data was found to predict valve performance with reasonable accuracy. This type of model had its advantages over the physically based model because it was easily implemented within standard simulation packages. Also, the valve manufacturer performs quality assurance tests to insure that the parameters of the experimentally based model are within allowable limits. Because of the stated advantages, the experimentally based model was selected to model the servovalve located within the hydraulic system of the 70-meter antenna.

II. Theoretical Model Development

A. Presentation of Mathematical Equations

Previous experimental investigations have shown that the dynamic response of a flow control servovalve can be approximated by the second-order equations [3,6,7,8]

$$\frac{\partial^2 Q_{c1}}{\partial t^2} = \omega_n^2 \Delta Q_{c1} - 2\xi\omega_n \frac{\partial Q_{c1}}{\partial t} \quad (1)$$

$$\frac{\partial^2 Q_{c2}}{\partial t^2} = \omega_n^2 \Delta Q_{c2} - 2\xi\omega_n \frac{\partial Q_{c2}}{\partial t} \quad (2)$$

where

$$\begin{aligned} \Delta Q_{c1} &= K_v U \sqrt{\frac{2|P_S - P_{c1}|}{\rho}} \\ &\times \operatorname{sgn}(P_S - P_{c1}) - Q_{c1} \quad U \geq 0 \end{aligned} \quad (3)$$

$$\begin{aligned} \Delta Q_{c1} &= K_v U \sqrt{\frac{2|P_{c1} - P_T|}{\rho}} \\ &\times \operatorname{sgn}(P_{c1} - P_T) - Q_{c1} \quad U < 0 \end{aligned} \quad (4)$$

$$\begin{aligned} \Delta Q_{c2} &= -K_v U \sqrt{\frac{2|P_{c2} - P_T|}{\rho}} \\ &\times \operatorname{sgn}(P_{c2} - P_T) - Q_{c2} \quad U \geq 0 \end{aligned} \quad (5)$$

$$\begin{aligned} \Delta Q_{c2} &= -K_v U \sqrt{\frac{2|P_S - P_{c2}|}{\rho}} \\ &\times \operatorname{sgn}(P_S - P_{c2}) - Q_{c2} \quad U < 0 \end{aligned} \quad (6)$$

$$U = i - i_{NB} + i_T + i_H + i_X$$

$$|i - i_{NB} + i_T + i_H + i_X| \leq i_{SAT} \quad (7)$$

$$U = i_{SAT} \operatorname{sgn}(i - i_{NB} + i_T + i_H + i_X)$$

$$|i - i_{NB} + i_T + i_H + i_X| > i_{SAT} \quad (8)$$

$$K_v = K_1 \quad |i - i_{NB} + i_T + i_H| \leq i_{cr} \quad (9)$$

$$K_v = K_2 \quad |i - i_{NB} + i_T + i_H| > i_{cr} \quad (10)$$

$$i_X = 0 \quad |i - i_{NB} + i_T + i_H| \leq i_{cr} \quad (11)$$

$$i_X = i_X^* \operatorname{sgn}(i - i_{NB} + i_T + i_H)$$

$$|i - i_{NB} + i_T + i_H| > i_{cr} \quad (12)$$

The above variables, to be described in greater detail later in this article, are defined as

i = input current to the valve

i_{cr} = critical current defining the boundary of the null flow region

i_H = servovalve hysteresis current

i_{NB} = servovalve null bias current

i_{SAT} = lowest current input at which the flow limit of the valve is reached

i_T = compensating current to model threshold

i_X = X-axis intercept current adjustment

i_X^* = X-axis intercept current relative to null

K_v = valve proportionality constant

i_X^* = X-axis intercept current relative to null

K_v = valve proportionality constant

K_1 = valve proportionality constant near null

K_2 = valve proportionality constant away from null

P_{c1} = pressure at control port no. 1

P_{c2} = pressure at control port no. 2

P_S = supply pressure

P_T = tank or return pressure

Q_{c1} = flow rate out of control port no. 1

Q_{c2} = flow rate out of control port no. 2

t = time

U = effective command signal

ρ = fluid density

ω_n = valve natural frequency

ξ = damping ratio

B. Pressure-Drop-to-Flow Relationship

A servovalve operates by metering fluid flow from the supply line to one control port and from the other control port to the return line by using a restriction with a pressure-drop-to-flow relationship equivalent to that of an orifice. Using the orifice model, the fluid flow rate from the supply line to the control ports and from the control ports to the return line varies with the square root of the pressure difference across the valve and as the square root of

the inverse of the fluid density. Equations (3) through (6) provide a mathematical description of the orifice behavior.

C. Flow Gain

Servovalves are manufactured to produce a linear flow gain with respect to input current for a given valve pressure drop. In practice the flow gain is modeled as a piecewise linear function. There are three linear regions of the flow curve as shown in Fig. 1. The three regions are referred to as the null region, the normal flow region, and the flow saturation region. Within the null region where the current input to the valve is $\pm i_{cr}$ of the null position, the valve exhibits a linear flow gain K_1 according to Eq. (9). The value of the flow gain K_1 is dependent upon the lap of the valve. Critically lapped valves exhibit a flow gain that is equivalent to the flow gain K_2 found in the region of normal flow while valves that are overlapped and underlapped have flow gains that are lower and higher than K_2 , respectively. The region of normal flow is the designed operating range of the valve where the rated flow gain K_2 occurs. Once the current input to the valve exceeds a certain level, flow saturation occurs. When saturation exists the flow gain decreases in a nonlinear fashion until the flow limit of the valve is reached. At the flow limit of the valve an increase in current input to the valve results in no increase of flow to the control port. The behavior of the flow gain in the flow saturation region is governed by such things as the stability of the supply pressure at high valve flow rates and the mechanical limitations placed on the valve spool. Since the supply pressure to the servovalve is fairly constant on the 70-meter antenna, the region of saturation is modeled as an abrupt change from the region of normal flow to the valve flow limit.

D. Null Bias

The null bias of a valve is the input current required to bring the valve to its null position under a certain set of operating conditions not including the effects of hysteresis. The null bias of the valve is modeled in Eq. (7) by the variable i_{NB} .

E. Threshold

Threshold is the increment of current input required to produce a change in valve flow rate when changing the direction of the applied current to the valve. A valve flow curve illustrating threshold is shown in Fig. 2. Threshold is mathematically modeled in Eq. (7) by the variable i_T and is computed once every sampling period that a current signal is sent to the servovalve by the following logic statements:

$$\text{If } i > i_k + \varepsilon \text{ then} \quad (13)$$

$$F_{k+1} = 1 \quad (14)$$

$$i_{k+1} = i \quad (15)$$

$$\text{else if } i < i_k - \varepsilon \text{ then} \quad (16)$$

$$F_{k+1} = -1 \quad (17)$$

$$i_{k+1} = i \quad (18)$$

else

$$F_{k+1} = F_k \quad (19)$$

$$i_{k+1} = i_k \quad (20)$$

end if

$$\text{If } F_{k+1} \times F_k = -1 \text{ then} \quad (21)$$

$$\text{if } i_k > i_A \text{ then} \quad (22)$$

$$i_A = i_k \quad (23)$$

$$i_B = i_A - i_T^* \quad (24)$$

$$\text{else if } i_k < i_B \text{ then} \quad (25)$$

$$i_B = i_k \quad (26)$$

$$i_A = i_B + i_T^* \quad (27)$$

end if

end if

$$\text{If } i > i_A \text{ then} \quad (28)$$

$$i_T = -\frac{1}{2}i_T^* \quad (29)$$

$$\text{else if } i < i_B \text{ then} \quad (30)$$

$$i_T = \frac{1}{2}i_T^* \quad (31)$$

else

$$i_T = i_A - i - \frac{1}{2}i_T^* \quad (32)$$

end if

$$k = k + 1 \quad (33)$$

where

F_k = logical variable that indicates the direction of application of the input current. The variable takes on the values of -1 for decreasing current and $+1$ for increasing current.

i = input current to the valve

i_A = upper endpoint of current range where flow does not change

i_B = lower endpoint of current range where flow does not change

i_k = input current to the valve at sample interval k

i_T = input current adjustment required to model threshold

i_T^* = threshold current

k = sample interval

ε = amplitude of input current signal noise

Equations (13) through (20) are designed to keep track of whether the input current is increasing or decreasing while taking into consideration the amplitude of the input signal noise. A change in the direction of current application is detected by Eq. (21), while Eqs. (22) through (27) are used to redefine the current region where invariant flow occurs, as shown in Fig. 3. The adjustment to the input current required to model the threshold phenomenon is computed using Eqs. (28) through (32). Equation (33) is used to increment the sampling interval index. The logical statements used to model threshold have been restated in flow diagram format in Fig. 4.

F. Hysteresis

During a single cycle of the input signal, a given flow rate is achieved twice: once as the signal increases and once as it decreases. Hysteresis is the difference in valve current inputs required to produce this flow rate. The magnitude of hysteresis increases with the amplitude at which the input signal is cycled, as shown in Fig. 5. When the current is cycled at low amplitudes, the hysteresis effects are contained within the threshold model. As the amplitude of the current cycle is increased, additional modeling is required to predict valve performance. During the operation of the hydraulic system of the 70-meter antenna, the servovalve will operate closely around a single input current as opposed to cycling through its full operating range. Under these conditions the threshold model is sufficient to describe the hysteresis effects of the valve. Because the slewing of the antenna operates over the full region of the valve, the flow curve of the valve may shift within the limits of the valve hysteresis depending upon the current signal history prior to antenna tracking. The shift of the flow gain curve due to hysteresis is modeled by the variable i_H in Eq. (7). The variable i_H is treated as a statistical variable that changes over extended periods of time and takes on values within a specific distribution.

III. Experimental Investigation

A. Overview

During the investigation of the performance characteristics of the Moog 72-163B servovalve at JPL, experimental programs were undertaken both at Moog under the

direction of Peter Hames and at Fluid Technologies, Inc., under the direction of Elizabeth Carrell. The results gathered during these investigations were used to identify the parameters of the theoretical model presented in the previous section and to estimate the accuracy of the model.

B. Frequency Response

The frequency response characteristics of the Moog 72-163B servovalve were determined through experiments conducted at Moog on January 27, 1988. It was concluded from an analysis of the results that the dynamic response of the servovalve is best approximated by a second-order model with a natural frequency of 55 hertz and a damping ratio of 0.8. Both the experimental frequency response and the second-order model approximations are presented in Figs. 6 and 7.

C. Hysteresis and Threshold

The hysteresis and threshold effects of the Moog 72-163B servovalve no. 110 were experimentally determined at Moog on April 19, 1988. The experimental results indicated a hysteresis of 0.1 milliamperes when the valve was cycled between negative and positive full flow. A threshold current of 0.01 milliamperes was also observed during the testing of the valve.

D. Steady State Flow Characteristics

An experimental program was conducted at Fluid Technologies, Inc., to determine the effect that the geometrical dimensions of the valve spool have upon servovalve flow characteristics. The program was conducted between December 31, 1990, and January 7, 1991. The spool dimensions along with the supply pressures for each of the experiments are shown in Fig. 8. Details of the experimental hardware and procedures are presented in [1,2,4,5].

E. Data Reduction

The valve flow constant K_v and the X-axis intercept of the flow curve were computed by performing a least squares fit of the experimental flow rate versus input current data using the computer program MATLAB. The data points contained in the negative, null, and positive flow regions of the servovalve were fit to a first-order polynomial separately in order to conform to the piecewise linear theoretical model. Some of the data points obtained at the higher current input magnitudes were removed prior to fitting the model to the data. The removal of the data points believed to be invalid was done because the valve was observed to saturate at flow rates near the maximum output capability of the pump supplying fluid to the servovalve. Therefore, it was suspected that the supply pres-

output capability of the pump supplying fluid to the servovalve. Therefore, it was suspected that the supply pressure to the valve was not constant at the larger input current magnitudes. There was no way of knowing whether the supply pressure was actually constant since the supply pressure was not monitored continuously during the experiments. Hence, some data points were removed where saturation of the valve was observed. The parameters that were identified from the least squares fit of the data are presented in Table 1, while a plot of both the experimental and theoretical valve flow curves are shown by the example in Fig. 9. It is evident from an examination of Fig. 9 that the experimental flow curves of the positive and negative flow regions did not necessarily match the flow curve of the null region from a continuity standpoint. This lack of continuity occurs because the data taken in the null region were obtained in a separate experimental run from the data taken in the normal flow region. Hence, the observed discontinuity is attributed to the effects of hysteresis and threshold. For this reason, the data taken in the null region were used only to identify the valve flow constant. The results presented in Table 1 indicate that the valve-flow-gain constant, K_v , is slightly different for the positive and negative flow regions contrary to expectation. This small difference is believed to be the result of the valve spool and bushing not being "exactly" symmetrical about null. When modeling the valve in computer simulation programs, an average of the valve flow constant in the positive and negative flow regions should be used to model the valve in the region of normal flow. The null bias of the valve was taken to be the midpoint of where the positive and negative flow curves intersect the X-axis. The value

of the X-axis intercept current, i_x^* , used in the servovalve model is taken to be one half the current magnitude between the locations where the positive and negative flow curves intersect the X-axis. The current value for the variable i_{cr} that defines the boundary of the null flow region is computed by finding the current at which the linear flow curve of the null region intersects the linear flow curve of the normal flow region. The servovalve model parameters for each of the experimental valves is presented in Table 2.

F. Summary of Results

Experimental methods were implemented to determine the parameters for the Moog 72-163B servovalve according to the servovalve model presented in Section II. The model parameter values that should be used when simulating the valve in the hydraulic system of the 70-meter antenna are presented in Tables 2 and 3.

IV. Conclusion

In order to more completely understand the dynamic behavior of the servo hydraulic system of the 70-meter antennas located in the DSN, a mathematical model for a servovalve was developed. Experimental data were used to identify the model parameters for the Moog 72-163B servovalves that are incorporated into the hydraulic system of the 70-meter antenna. The servovalve performance predicted by the theoretical servovalve model was found to be in excellent agreement with the experimental data.

Acknowledgments

The author thanks Fredrick J. Menninger and Ben A. Parvin for their support during the experimental program.

References

- [1] E.-A. Carrell, "Investigation of Mechanical Wear Effects on Servovalve Performance," Master of Science thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, June 1991.
- [2] A. Khalil, *Servovalve Performance Assessment—Final Report*, Report no. VT90108, Fluid Technologies, Inc., Stillwater, Oklahoma, March 1991.
- [3] H. E. Merritt, *Hydraulic Control Systems*, New York: John Wiley & Sons, 1967.
- [4] *Catalog 72-890A*, Moog Inc., East Aurora, New York.
- [5] *Electrohydraulic Servovalve 72 Series Service Manual*, Moog Inc., East Aurora, New York, 1990.
- [6] W. J. Thayer, *Specifications Standards for Electrohydraulic Flow Control Servovalves*, Technical Bulletin 117, Moog Inc., East Aurora, New York, 1962.
- [7] W. J. Thayer, *Transfer Functions For Moog Servovalves*, Technical Bulletin 103, Moog Inc., East Aurora, New York, 1962.
- [8] J. Watton, *Fluid Power Systems Modeling, Simulation, Analog, and Microcomputer Control*, New York: Prentice Hall, 1989.

Table 1. Servovalve parameters.

Valve ID	Negative flow region			Null flow region			Positive flow region		
	K_v , m ² /mA	I_x , mA	σ^2 , (m ³ /sec) ²	K_v , m ² /mA	I_x , mA	σ^2 , (m ³ /sec) ²	K_v , m ² /mA	I_x , mA	σ^2 , (m ³ /sec) ²
A-1	1.08E-06	-5.43E-01	1.89E-10	3.90E-07	-3.94E-01	8.30E-10	1.18E-06	3.01E-01	1.40E-10
A-2	1.05E-06	-3.73E-01	2.45E-10	2.83E-06	2.43E-01	6.81E-10	1.13E-06	6.77E-01	6.86E-11
A-3	1.08E-06	-1.46E-01	2.53E-10	1.99E-06	5.85E-01	5.63E-10	1.07E-06	8.87E-01	1.03E-09
A-4	9.99E-07	-6.17E-01	5.40E-11	1.03E-06	8.86E-02	1.71E-09	1.04E-06	5.29E-01	8.30E-11
B-1	1.15E-06	-1.42E+00	4.51E-11	7.90E-08	-8.68E-01	5.60E-12	1.20E-06	2.80E-01	1.96E-10
B-2	1.06E-06	-4.53E-01	7.88E-11	3.38E-07	-1.32E-01	2.73E-11	1.11E-06	5.74E-01	1.06E-10
B-3	1.17E-06	-4.29E-01	7.52E-11	1.45E-07	3.28E-02	3.12E-12	1.23E-06	4.36E-01	5.55E-11
B-4	1.02E-06	-3.16E-01	8.06E-11	8.72E-07	4.21E-02	2.52E-10	1.05E-06	2.18E-01	9.37E-11
B-5	1.13E-06	-3.22E-01	1.57E-10	1.01E-06	-3.63E-03	1.80E-10	1.17E-06	1.15E-01	1.43E-10
C-1	1.05E-06	-6.60E-01	5.35E-11	3.27E-07	-2.10E-01	9.29E-11	1.15E-06	4.32E-01	7.28E-11
C-2	1.17E-06	-8.48E-01	1.60E-10	5.11E-07	-6.99E-01	2.27E-11	1.26E-06	-2.07E-01	6.94E-11
C-3	1.04E-06	-7.54E-01	8.08E-11	1.13E-07	-8.64E-02	5.00E-12	1.10E-06	6.58E-01	1.34E-10
C-4	1.02E-06	-7.96E-01	4.36E-11	3.00E-09	3.24E-01	5.00E-13	1.06E-06	1.15E+00	1.46E-10
D-1	1.46E-06	7.02E-03	4.41E-10	1.08E-06	4.12E-02	4.91E-11	1.53E-06	3.10E-01	2.76E-10
D-2	1.23E-06	-5.14E-01	1.95E-10	1.14E-06	3.02E-01	1.32E-10	1.38E-06	6.73E-01	9.01E-10

K_v = valve proportionality constant.

I_x = X-axis intercept.

σ^2 = variance of the residuals.

Table 2. Servovalve model parameters.

Valve ID	K_1 , m ² /mA	K_2 , m ² /mA	i_X^* , mA	i_{cr} , mA	i_{NB} , mA
A-1	3.89E-07	1.13E-07	4.22E-01	6.45E-01	-1.22E-01
A-2	2.83E-06	1.09E-06	5.25E-01	-3.27E-01	1.52E-01
A-3	1.99E-06	1.07E-06	5.17E-01	-6.08E-01	3.71E-01
A-4	1.03E-06	1.02E-06	5.73E-01	-6.85E+01	-4.39E-02
B-1	7.91E-08	1.18E-06	8.52E-01	9.13E-01	-5.71E-01
B-2	3.38E-07	1.09E-06	5.14E-01	7.45E-01	6.04E-02
B-3	1.45E-07	1.20E-06	4.32E-01	4.92E-01	3.08E-03
B-4	8.72E-07	1.04E-06	2.67E-01	1.69E+00	-4.89E-02
B-5	1.01E-06	1.15E-06	2.19E-01	1.80E+00	-1.04E-01
C-1	3.27E-07	1.10E-06	5.46E-01	7.76E-01	-1.14E-01
C-2	5.11E-07	1.22E-06	3.21E-01	5.53E-01	-5.27E-01
C-3	1.13E-07	1.07E-06	7.06E-01	7.89E-01	-4.80E-02
C-4	2.85E-09	1.04E-06	9.72E-01	9.75E-01	1.77E-01
D-1	1.08E-06	1.49E-06	1.51E-01	5.48E-01	1.58E-01
D-2	1.14E-06	1.31E-06	5.93E-01	4.65E+00	7.96E-02

K_1 = valve proportionality constant in the null flow region.

K_2 = valve proportionality constant in the normal flow region.

i_x^* = X-axis intercept relative to null.

i_{cr} = critical current defining the null flow region.

i_{NB} = null bias.

Table 3. Servovalve model parameters.

Parameter	Value
i_H , mA	± 0.1
i_{SAT} , mA	20 (rated current)
i_T^* , mA	0.01
P_S , bar	172.4
P_T , psi	0
ρ , kg/m ³	854
ω_n , Hz	55
ξ	0.8

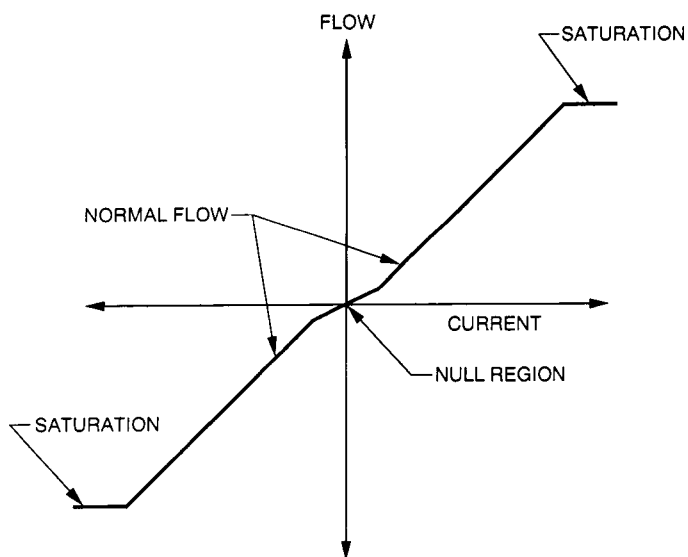


Fig. 1. Servovalve flow curve at a constant pressure drop.

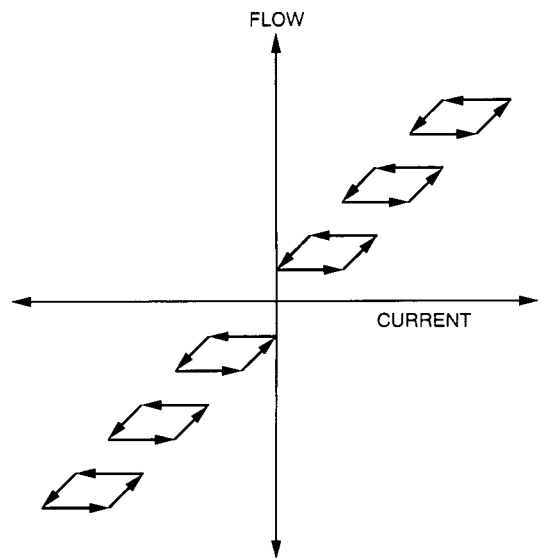


Fig. 2. Threshold loops.

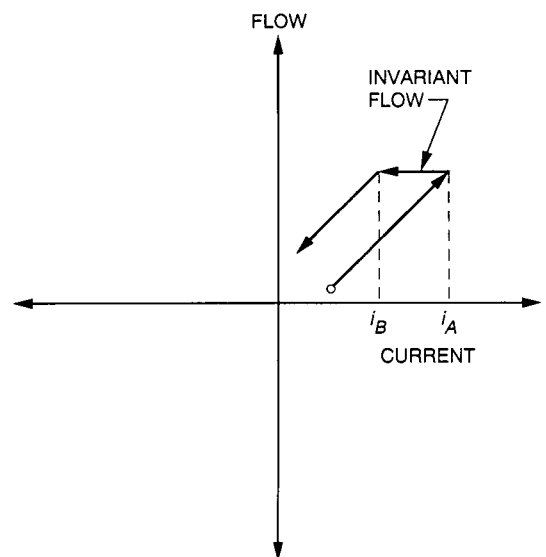


Fig. 3. Invariant flow region.

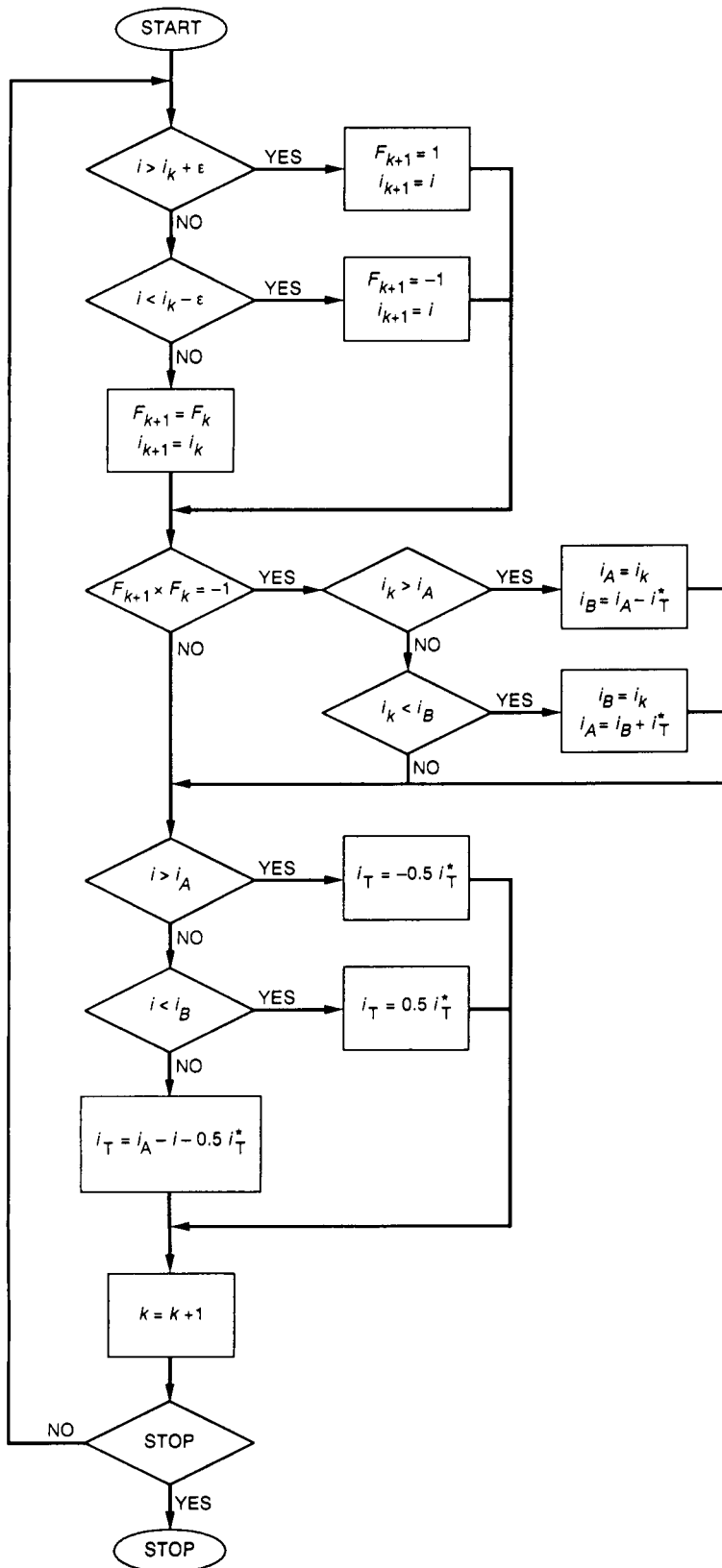


Fig. 4. Threshold model.

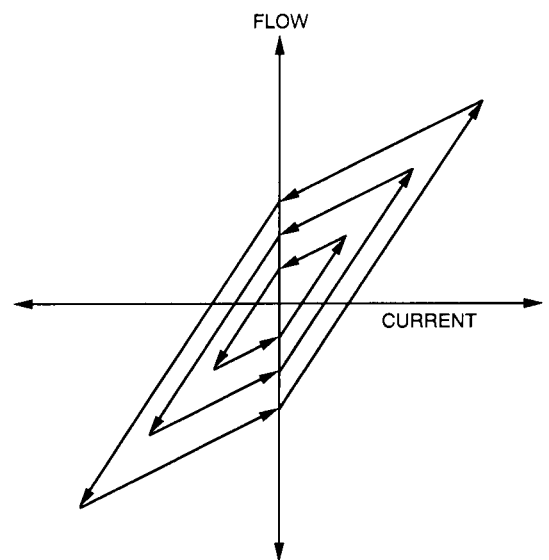


Fig. 5. Servovalve hysteresis.

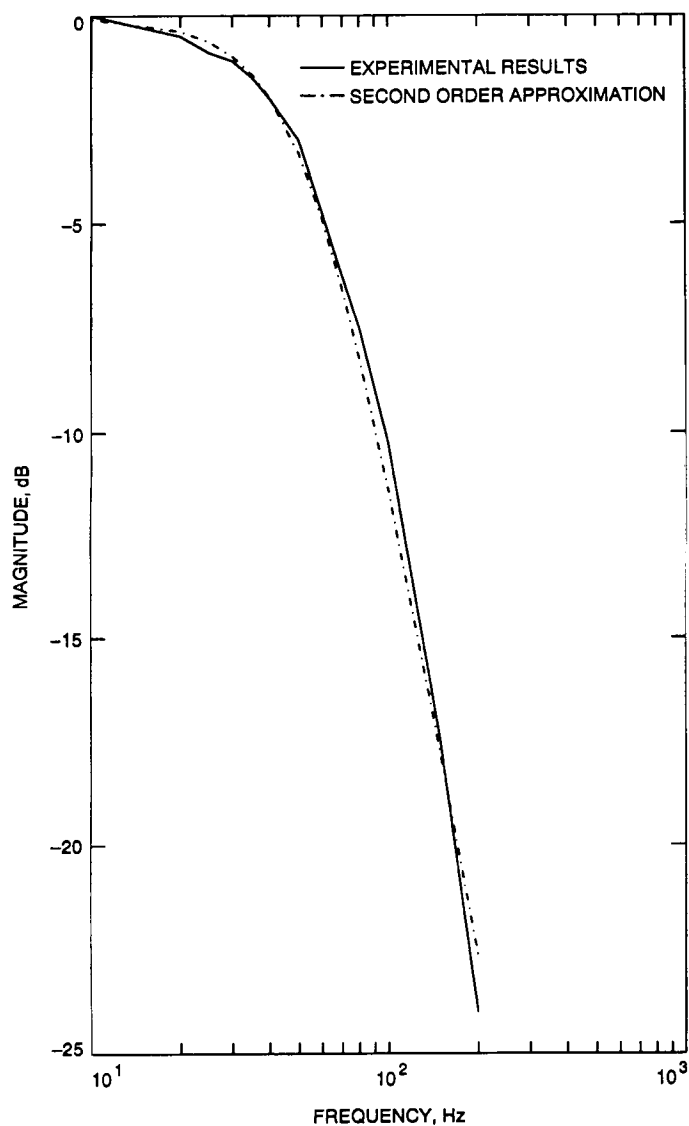


Fig. 6. Magnitude frequency response of Moog 72-163B servo-valve; supply pressure = 172.4 bar.

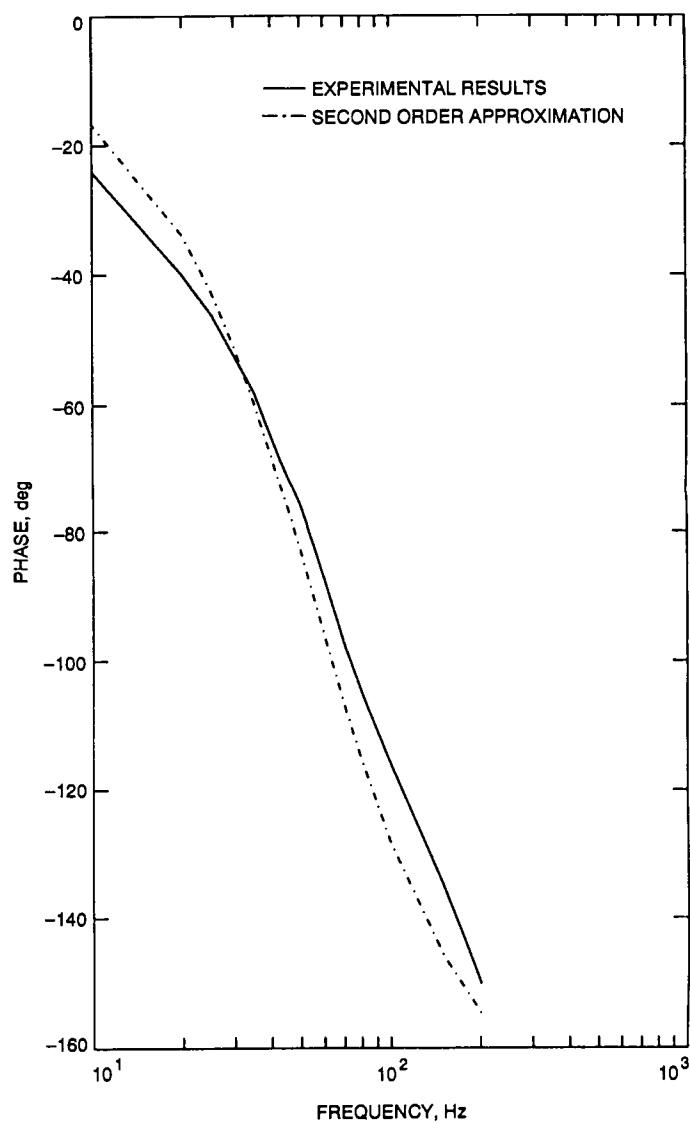
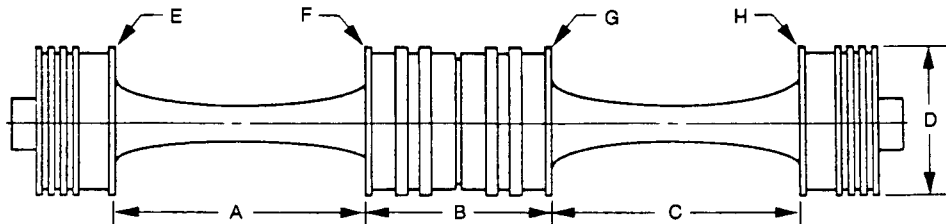


Fig. 7. Phase angle frequency response of Moog 72-163B servo-valve; supply pressure = 172.4 bar.



VALVE ID	SPOOL DIMENSIONS								PRESSURE (PSI)
	A (in.)	B (in.)	C (in.)	D (in.)	E (mm)	F (mm)	G (mm)	H (mm)	
A-1	0.999	1.003	1.001	0.74975	0.03	0.03	0.03	0.03	2500
A-2	"	"	"	"	0.08	0.09	0.09	0.09	2500
A-3	"	"	"	"	0.15	0.17	0.16	0.15	2500
A-4	"	"	"	"	0.32	0.35	0.34	0.33	2500
B-1	0.993	1.005	0.9965	0.7498	0.05	0.01	0.04	0.03	2500
B-2	0.9963	1.0017	0.9986	"	0.04	0.02	0.03	0.02	2500
B-3	REPEAT OF VALVE B-2								1800
B-4	1	0.9976	0.9975	"	0.03	0.03	0.02	0.02	2500
B-5	REPEAT OF VALVE B-4								1800
C-1	1.0005	0.9975	0.9995	0.7499	0.03	0.02	0.04	0.04	2500
C-2	REPEAT OF VALVE C-1								1800
C-3	"	"	"	0.7495	0.02	0.02	0.03	0.03	2500
C-4	"	"	"	0.7491	0.01	0.01	0.01	0.01	2500
C-5	REPEAT OF VALVE C-4								1800
D-1	0.9975	0.997	0.998	0.7499	0.05	0.05	0.06	0.14	2500
D-2	0.005 AXIAL SCRATCH EQUALLY SPACED THREE PLACES								2500
D-3	0.010 AXIAL SCRATCH EQUALLY SPACED THREE PLACES								2500

Fig. 8. Servovalve spool dimensions and locations.

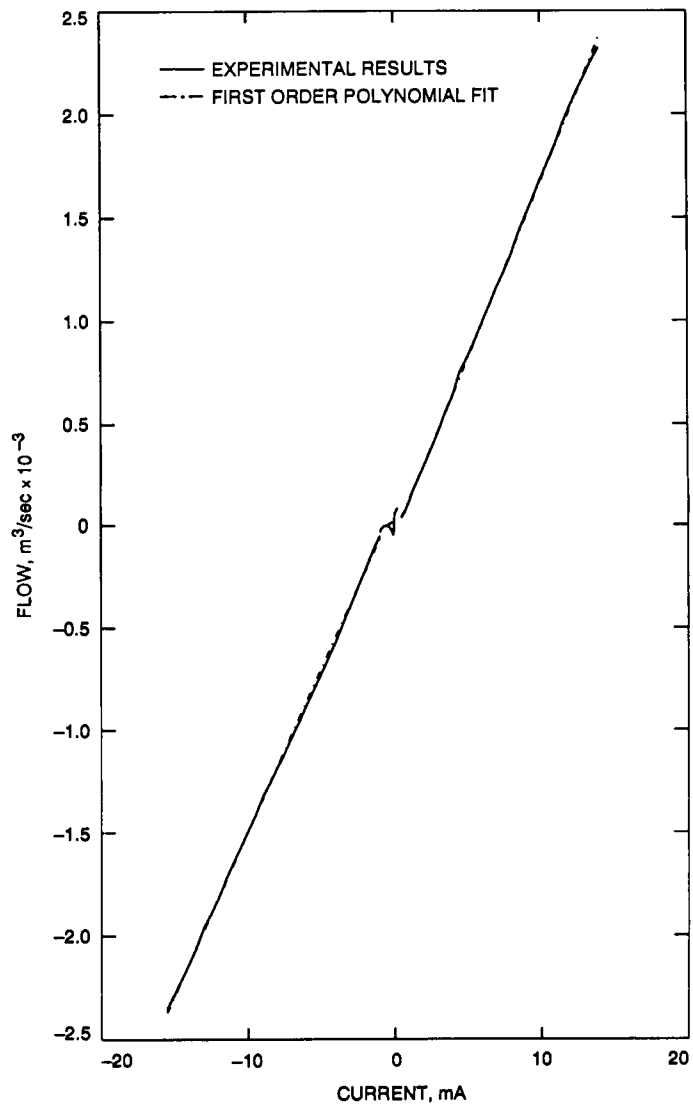


Fig. 9. Flow gain curve of servovalve B-3; supply pressure = 124.1 bar.

48-35
79878
N92-24819

JJ574450

A Microwave Holography Methodology for Diagnostics and Performance Improvement for Large Reflector Antennas

D. J. Rochblatt

Ground Antennas and Facilities Engineering Section

Microwave holography has proven to be a powerful technique for various evaluations, diagnostics, and RF performance improvements for large reflector antennas. This technique utilizes the Fourier transform relation between the complex far-field radiation pattern of an antenna and the complex aperture field distribution. Resulting aperture phase and amplitude distribution data can be used to precisely characterize various crucial performance parameters, including panel alignment, panel shaping, subreflector position, antenna aperture illumination, directivity at various frequencies, and gravity deformation effects.

The methodology of the data processing presented in this article was developed at JPL and has been successfully applied to the DSN 34-m beam waveguide antennas. A companion article in this issue (Rochblatt and Seidel) describes the application of this technology to the DSS-13 antenna. The antenna performance was improved at all operating frequencies (wide-bandwidth improvement) by reducing the main reflector mechanical surface rms error to 0.43 mm. At Ka-band (32-GHz), the estimated improvement is 4.1 dB, resulting in an aperture efficiency of 52 percent. The performance improvement was verified by efficiency measurements and additional holographic measurements.

I. Introduction

Microwave holography is a measurement technique that has now been successfully applied to improve the performance of many types of reflector and array antennas throughout the world [1-7].¹ The raw data (the observable) for this technique is the complex far-field pattern

of the antenna under test. For large reflector antennas, geostationary satellites are commonly used as convenient far-field signal sources. The beacon continuous wave (CW) signal from a geostationary satellite is usually used with a narrow-bandwidth receiver, while a transponder signal is usually the preferred choice for a wide-bandwidth receiver. Figure 1 presents a typical narrow-bandwidth receiver system architecture based on an HP 8510B analyzer and an external phase-locked loop (PLL) (with a variable PLL bandwidth of 1 to 20 Hz). The reference antenna is

¹ B. L. Seidel and D. J. Rochblatt, *DSS-13 Beam Waveguide Antenna Phase I Final Report*, JPL D-8451 (internal document), Jet Propulsion Laboratory, Pasadena, California, pp. 4.1-4.23, May 15, 1991.

ORIGINAL CONTAINS
COLOR ILLUSTRATIONS

needed as a phase reference for the measurement, as well as to keep the narrow-bandwidth receiver in phase lock with the carrier [8,9].

This article provides a detailed analysis of the unique methodology and data processing developed at JPL by the author.

II. Mathematical Algorithms

The mathematical relationship between an antenna far-field radiation pattern (T) and the antenna surface-induced current distribution (J) is given by the exact radiation integral [10]

$$\vec{T}(u, v) = \int_s \int \tilde{J}(x', y') \exp^{jkz'} [\exp^{-jkz'(1-\cos\theta)}] \times \exp^{j(ux'+vy')} dx' dy' \quad (1)$$

where

$$\begin{aligned} z(x', y') &= \text{the surface } s \\ (u, v) &= \text{directional cosine space} \\ \theta &= \text{observation angle} \end{aligned}$$

For small observation angles ($\theta \cong 0$ deg) of the far-field pattern, this expression becomes a Fourier transform relationship

$$\vec{T}(u, v) = \int_s \int \tilde{J}(x', y') \exp^{jkz'} \exp^{j(u x' + v y')} dx' dy' \quad (2)$$

To derive the residual surface error, geometrical optics-ray tracing is used to relate the small normal error, ϵ , directly to an aperture phase error in a main reflector paraboloid geometry (Fig. 2). In this small error approximation, it is assumed that the aperture phase error is entirely due to the projected surface errors. In addition, the three-dimensional structure of the surface error is not recovered; rather, an axial component equal to the average error over the resolution cell size is recovered. The normal error can then be computed using the paraboloid geometry. In practical application, when many resolution cells resolve a single panel, high accuracy (typically $\lambda/250$ at 12 GHz) can be achieved in determining panel setting screw adjustments. It can be computed from the geometry in Fig. 2

$$1/2\Delta PL \equiv 1/2[P'P + PQ]$$

$$= 1/2 \left[\frac{\epsilon}{\cos\phi} + \frac{\epsilon \cos 2\phi}{\cos\phi} \right] = \epsilon \cos\phi \quad (3)$$

$$\text{Phase}(\Delta PL) = \frac{4\pi}{\lambda} \epsilon \cos\phi \quad (4)$$

and

$$\cos\phi = \frac{1}{\sqrt{1 + \frac{Z}{F}}} \quad (5)$$

and for a paraboloid:

$$\cos\phi = \frac{1}{\sqrt{1 + \frac{X^2 + Y^2}{4F^2}}} \quad (6)$$

Allowing for the removal of a constant phase term and substituting Eq. (4) into Eq. (2) yields:

$$T(u, v) = \exp^{-j2kF} \int_s \int [J(x', y')] \exp^{j4\pi\epsilon/\lambda \cos\phi} \exp^{jkz'} \times \exp^{j(u x' + v y')} dx' dy' \quad (7)$$

where $[...] \equiv J_e$ (equivalent current distribution).

For the processing of sampled data, the associated discrete Fourier transform (DFT) is utilized:

$$T(p\Delta u, q\Delta v) = sx sy \sum_{n=-N1/2}^{N1/2-1} \sum_{m=-N2/2}^{N2/2-1} J(nsx, msy) \times \exp^{j2\pi \left(\frac{np}{N1} + \frac{mq}{N2} \right)} \quad (8)$$

where

$N1, N2$ = the measured data array size

sx, sy = sampling intervals in the aperture coordinates

n, m, p, q = integers indexing the discrete samples

$\Delta u, \Delta v$ = sampling intervals in u, v far-field space

Since the magnitude of the far-field pattern is essentially bounded, the fast Fourier transform (FFT) is usually used for computation, and is symbolized here by (F) . Solving for the residual normal surface error and substituting Eq. (6) yields

$$\varepsilon(x, y) = \frac{\lambda}{4\pi} \sqrt{1 + \frac{x^2 + y^2}{4F^2}} \text{Phase} \{ \exp^{j2kF} F^{-1}[T(u, v)] \} \quad (9)$$

The spatial resolution in the final holographic maps is defined at the 50-percent level of the Fourier transform of the truncation function [7,11]. Rectangular and triangular windowing are commonly used, with competing advantages to both. For rectangular windowing, the lateral resolution is derived [7,11]

$$\delta = \frac{D}{kN} \quad (10)$$

where

D = main reflector diameter

N = the square root of the total number of data points

k = sampling factor, usually $0.5 < k < 1.0$

The lateral resolution is inversely proportional to the number of sidelobes measured. For the DSN 70-m antennas, the required resolution of 0.4 m is achieved with a data array size of 197×197 . For the DSN 34-m antennas, the required resolution of 0.32 m is achieved with a data array size of 127×127 [12].

Figure 3 shows the complex Fourier transform relationship between the far-field and the aperture functions for a 34-m antenna (DSS 13). The far-field amplitude and phase, Figs. 3(a) and 3(b), respectively, are measured on rectangular coordinates of 127×127 with sampling intervals of 34.8 mdeg (the sampling factor is 0.84). Figures 3(c) and 3(d) show the aperture amplitude and surface error function with a lateral resolution of 0.32 m.

The accuracy in each resolution cell of the final holographic maps was formulated from a simulation study [9], which agrees well with analytically derived expressions [2]. In this study [9], Eq. (11) is derived

$$\sigma \cong 0.082 \frac{\lambda D}{\delta SNR} \quad (11)$$

where

σ = standard deviation (accuracy) in recovering the mean position of a resolution cell

λ = wavelength

SNR = beam peak voltage signal-to-noise ratio in the test (antenna) channel

The accuracy across the holographic map varies with the aperture amplitude illumination. The accuracy is better at the center of the dish and gradually becomes worse toward the edge, where the illumination falls off rapidly [9]. For a uniformly illuminated antenna design, the accuracy remains relatively constant, becoming quickly worse just at the edge, where the illumination falls off rapidly. Typically, accuracies of 0.05 mm to 0.10 mm were achieved with a corresponding SNR of 73 dB to 60 dB. Equation (11) also indicates that the accuracy is inversely proportional to the resolution. This is due to the larger averaging area available at the larger resolution cell, as expected.

The resulting aperture phase function (the dirty map) needs to be corrected for modulo- 2π phase errors. These occur due to the small measurement wavelength and a large phase error that is partially due to pointing and subreflector position errors. Figures 4 and 5 are examples of aperture phase maps before and after correction of modulo- 2π phase errors, respectively.

The next phase cleanup algorithm removes from the aperture function phase error components that are due to pointing and subreflector position errors. This can be done by fitting polynomials to the aperture phase data whose orders are equal to the expected order of the errors. Either the Zernike [13] or the modified Jacobi polynomials [14] can be used. Determination of pointing error-linear term, astigmatism, focus shift, coma, and spherical aberration are readily identified. The orthogonality properties of the Jacobi polynomial allow solving for its coefficients by

$$C_m^n = \frac{\zeta_n}{2\pi} \int_0^1 \int_0^{2\pi} \varepsilon(s', \phi') \cos(n\phi') F_m^n(s') s' ds' d\phi' \quad (12)$$

$$D_m^n = \frac{\zeta_n}{2\pi} \int_0^1 \int_0^{2\pi} \varepsilon(s', \phi') \sin(n\phi') F_m^n(s') s' ds' d\phi' \quad (13)$$

where

$\varepsilon(s', \phi')$ = surface error function

ζ_n = Neumann number

In addition, the above polynomials can be used to separate systematic (low-frequency) errors from random (high-frequency) error components by filtering (Fig. 6) [6].

Another technique for removing phase errors due to pointing and subreflector position errors has been developed via the global best fit paraboloid. Here, the entire data set is weighted-least-squares fitted to the paraboloid, permitting six degrees of freedom in the model, three vertex translations, two rotations, and a focal length change.

Minimization is performed on the sum squares of the residual path length changes S , given by

$$S = \sum_{i=1}^{N^2} \Gamma (\Delta PL_i)^2 A_i \quad (14)$$

where

Γ = support domain masking operator

ΔPL_i = path length change

A_i = amplitude weighting factor

The minimum for S is found by solving the six partial differential equations simultaneously

$$\left. \begin{aligned} \frac{\partial S}{\partial x_0} &= 2 \sum_{i=1}^{N^2} \Gamma \frac{\partial \Delta PL_i}{\partial x_0} \Delta PL_i A_i = 0 \\ \frac{\partial S}{\partial y_0} &= 2 \sum_{i=1}^{N^2} \Gamma \frac{\partial \Delta PL_i}{\partial y_0} \Delta PL_i A_i = 0 \\ \frac{\partial S}{\partial z_0} &= 2 \sum_{i=1}^{N^2} \Gamma \frac{\partial \Delta PL_i}{\partial z_0} \Delta PL_i A_i = 0 \\ \frac{\partial S}{\partial \alpha} &= 2 \sum_{i=1}^{N^2} \Gamma \frac{\partial \Delta PL_i}{\partial \alpha} \Delta PL_i A_i = 0 \\ \frac{\partial S}{\partial \beta} &= 2 \sum_{i=1}^{N^2} \Gamma \frac{\partial \Delta PL_i}{\partial \beta} \Delta PL_i A_i = 0 \\ \frac{\partial S}{\partial K} &= 2 \sum_{i=1}^{N^2} \Gamma \frac{\partial \Delta PL_i}{\partial K} \Delta PL_i A_i = 0 \end{aligned} \right\} \quad (15)$$

where

$$\left. \begin{matrix} X_0 \\ Y_0 \\ Z_0 \end{matrix} \right\} = \text{vertex coordinates translation}$$

α = rotation about the x -axis

β = rotation about the y -axis

and

$$K = \frac{1}{4} \left(\frac{1}{F} - \frac{1}{F'} \right) \quad (16)$$

where

F = focal length of the original paraboloid

F' = focal length of the best-fit paraboloid

The antenna surface axial rms error is then computed with respect to the position of the fitting paraboloid

$$RMS = \frac{\sqrt{\sum_{i=1}^N \Gamma (1/2 \Delta PL_i)^2}}{\sum_{i=1}^{N^2} \Gamma} \quad (17)$$

The subreflector position error is found via the six parameters computed above. Figure 7 presents a large linear phase error term due to pointing errors where errors due to subreflector position are minimal.

It is correct to apply the best-fit paraboloid algorithm to either the conventional Cassegrain paraboloid-hyperboloid or dual-shaped reflector systems, even though the latter do not use a paraboloid as the main reflector. Either design is a plane-wave-to-point source transformer, differing only in the field intensity distribution.

The resultant aperture function at the end of this process is defined here as the effective map since it includes all phase effects that are contributing to the antenna performance [15]. These frequency-dependent effects include the subreflector scattered (frequency-dependent) feed phase function and struts' diffraction effects. Removal of the feed phase function and subreflector support structure diffraction effects results in a frequency independent map, which is defined here as the mechanical map.

By deriving panel adjustment based on the effective map, the surface shape will conjugate to the phase errors, optimizing the performance of the antenna at a single frequency equal to the measurement frequency, while degrading the performance at all other frequencies. For antennas operating at a single frequency, this procedure is advantageous. However, many antennas (including the DSN's)

operate at several different frequencies and require wide-bandwidth performance. For these antennas, the mechanical map must be used to derive panel setting information.

Figure 8 is a feed phase function derived for the DSS-13 antenna, a 34-m beam-waveguide antenna at 12.198 GHz, using Geometrical Theory of Diffraction (GTD). Figure 9 is an effective map derived for the same antenna at 12.198 GHz. Subtracting the feed phase function from the effective map and masking the struts' diffraction effects resulted in the mechanical map presented in Fig. 10. Successful removal of the frequency-dependent feed phase function must result in reduction of the rms error for the mechanical map. This is due to the uniqueness of such a solution. As shown above, the rms decreased by 27.5 percent, from 0.98 mm to 0.71 mm.

Panel setting information is derived by sorting all the data points within each panel and performing a rigid-body least-squares fit. The algorithm allows for one translation and two rotations (S^k , α^k , and β^k), hence, a rigid body motion (Fig. 11). For each panel and its associated n data points, the motion parameters are solved via Eq. (18)

$$\begin{bmatrix} \sum_{i=1}^n \cos^2(\gamma_i) & \sum_{i=1}^n d_i \cos^2(\gamma_i) & -\sum_{i=1}^n e_i \cos(\gamma_i) \\ \sum_{i=1}^n d_i \cos^2(\gamma_i) & \sum_{i=1}^n d_i^2 \cos^2(\gamma_i) & -\sum_{i=1}^n e_i d_i \cos(\gamma_i) \\ -\sum_{i=1}^n e_i \cos(\gamma_i) & -\sum_{i=1}^n d_i e_i \cos(\gamma_i) & \sum_{i=1}^n e_i^2 \end{bmatrix} \times \begin{bmatrix} S^k \\ \alpha^k \\ \beta^k \end{bmatrix} = \begin{bmatrix} -\sum_{i=1}^n \epsilon_i \cos^2(\gamma_i) \\ \sum_{i=1}^n \epsilon_i d_i \cos^2(\gamma_i) \\ \sum_{i=1}^n \epsilon_i e_i \cos^2(\gamma_i) \end{bmatrix} \quad (18)$$

This mathematical process increases the accuracy in determining the screw adjustment correction by a factor of \sqrt{n} , since all n data points within each panel are used to derive each screw adjustment. The three parameters from the least squares fit above are also used to recompute the predicted surface errors within each panel that will result after panel setting. Figures 12, 13, and 14 are, respectively, the before, predicted, and after mechanical maps derived during holographic measurements of DSS 13 and applying the above algorithms. The antenna was found to have an rms surface error of 0.88 mm, which was reduced to 0.45 mm. Figure 13, the predicted surface map (rms = 0.36 mm), indicates that panels in rings 8 and 9 were deformed, which was then confirmed in Fig. 14, the derived surface after panel setting.

The Appendix by S. Stewart, D. Rochblatt, and B. Seidel of the companion article in this issue describes in detail the cause for these deformed features due to using DSS-15 main reflector panels on the DSS-13 beam-waveguide (BWG) antenna. The reduction in rms from 0.88 mm to 0.45 mm was achieved within one session of panel setting, which limited the screw adjustments to the nearest 1/8 of a turn.² Figure 15 is a sample printout of the panel-setting screw-adjustment listing produced by this software package.

Another significant feature of the mechanical map is that surface tolerance efficiency can be computed at frequencies other than the measured one. This is possible since the aperture phase function has been cleaned of all effects that are frequency dependent. Surface tolerance efficiency can be computed by integrating the entire aperture function. In this computation, it is assumed that the aperture amplitude function is also frequency independent. The introduced error in this assumption is thus very small.

$$\eta(K)_{\text{surface}} = 20 \log_{10} \frac{\sqrt{\sum_{i=1}^{N^2} \Gamma 10^{ampdb_i/20} \cos \left[\phi_{m_i} \left(\frac{\lambda_m}{\lambda_k} \right) \right]^2 + \sum_{i=1}^{N^2} \Gamma 10^{ampdb_i/20} \sin \left[\phi_{m_i} \left(\frac{\lambda_m}{\lambda_k} \right) \right]^2}}{\sum_{i=1}^{N^2} \Gamma 10^{ampdb_i/20}} \quad (19)$$

where

ϕ_{m_i} = phase value at the measured frequency

sub k = any other frequency

An effective surface error ϵ can be computed by equating the above expression to that of Ruze [11,16]

²ibid.

$$\eta = \exp -(4\pi\epsilon/\lambda)^2 \quad (20)$$

Figure 16 shows the estimated surface loss versus frequencies of DSS 13 before and after panel setting, indicating a performance improvement of 4.1 dB at 32 GHz, which resulted in an aperture efficiency of 52 percent. Figure 17 shows the measured efficiency of the DSS-13 antenna before and after panel setting as a function of elevation angle at X-band (8.45 GHz). As shown in Fig. 17, the peak antenna gain increased by 0.21 dB and moved from the anti-symmetric 57-deg elevation to the desired 45-deg elevation that was specified as the required antenna rigging angle.

Figure 18 is a gravity deformation of DSS 13 between the elevation angles of 12.7 deg and 46.5 deg. Gravity

deformation maps are highly important in deriving accurate structural deformation models for large reflector antennas. It is observed from this image (Fig. 18), that the bypass shroud on the right side of the antenna is exerting a force in the outer axial direction, creating the observed signature. It was found that the disturbing shroud forces are attributed to eccentricity of the shroud axes of rotation with respect to the antenna elevation axis.³ The eccentricity is also confirmed by dial-indicator measurements. Figure 19 presents the major software components of the microwave holography methodology in a block diagram.

³ R. Levy, *DSS-13 Antenna Structure Measurements and Evaluation*, JPL D-8947 (internal document), Jet Propulsion Laboratory, Pasadena, California, October 1, 1991.

Acknowledgments

The author would like to thank Dan Bathker for his numerous helpful technical discussions and contributions. In addition, the author wishes to thank Boris Seidel and Manuel Franco from JPL and Paul Wright and Eric Schoessow from Eikontech, Ltd., for helping with the holographic measurements at DSS 13.

References

- [1] A. P. Anderson, J. C. Bennett, A. J. T. Whitaker, and M. P. Godwin, "Measurement and optimization of a large reflector antenna by microwave holography," *Proc. Int. Conf. on Ant. and Prop.*, IEE Publication no. 169, pp. 128-131, London, England, 1978.
- [2] P. F. Scott and M. Ryle, "A rapid method for measuring the figure of a radio telescope reflector," *Mon. Not. Roy. Astr. Soc.*, vol. 178, pp. 539-545, 1977.
- [3] J. C. Bennet, A. P. Anderson, P. A. McInnes, and A. J. T. Whitaker, "Microwave holographic metrology of large reflector antennas," *IEEE Trans. Antennas and Propagat.*, vol. AP-24, no. 3, pp. 295-303, May 1976.
- [4] M. P. Godwin, A. J. T. Whitaker, and A. P. Anderson, "Microwave diagnostics of the Chilbolton 25m antenna using OTS satellite," *Proc. Inst. Elec. Eng. Int. Conf.*, York, England, pp. 232-236, 1981.
- [5] C. E. Mayer, J. H. Davis, W. L. Peters, and W. J. Vogel, "A holographic surface measurement of the Texas 4.9-meter antenna at 86 GHz," *IEEE Trans. Instru. Meas.*, vol. IM-32, no. 1, pp. 102-109, March 1983.
- [6] D. J. Rochblatt, "Holographic measurements of 34 and 70m DSN antennas," *Antenna optics metrology workshop W.2*, presented at the URSI/IEEE-AP-S symposium, London, Ontario, Canada, June 28, 1991.

- [7] D. J. Rochblatt and B. L. Seidel, "DSN Microwave Antenna Holography," *TDA Progress Report 42-76*, vol. October–December 1983, Jet Propulsion Laboratory, Pasadena, California, pp. 27–42, February 15, 1984.
- [8] D. J. Rochblatt, "System Analysis for DSN Microwave Antenna Holography," *TDA Progress Report 42-97*, vol. January–March 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 132–157, May 15, 1989.
- [9] D. J. Rochblatt and Y. Rahmat-Samii, "Effects of Measurement Errors on Microwave Antenna Holography," *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 7, pp. 933–942, July 1991.
- [10] Y. Rahmat-Samii, "Surface diagnosis of large reflector antennas using microwave metrology—An iterative approach," *Radio Science*, vol. 19, no. 5, pp. 1205–1217, September–October 1984.
- [11] C. E. Mayer, J. H. Davis, and H. D. Foltz, "Texas 5-m antenna aperture efficiency doubled from 230–300 GHz with error compensating secondary," *IEEE Trans. Antennas and Propagat.*, vol. AP-39, no. 3, pp. 309–317, March 1991.
- [12] D. J. Rochblatt, "System requirements for NASA Deep Space Network microwave antenna holography," paper presented at International Symposium on Antenna and Propagation, Tokyo, Japan, August 22–25, 1989.
- [13] M. Born and E. Wolf, *Principles of Optics*, Oxford: Pergamon Press, pp. 464–468, 1965.
- [14] V. Galindo-Israel and R. Mittra, "A new series representation integral with application to reflector antennas," *IEEE Trans. Antennas Propagat.*, vol. AP-25, pp. 631–635, September 1977 (correction in *IEEE Trans. Antennas Propagat.*, vol. AP-26, p. 628, July 1978).
- [15] D. J. Rochblatt, Y. Rahmat-Samii, and J. H. Mumford, "DSN Microwave Antenna Holography Part II: Data Processing and Display of High Resolution Effective Maps," *TDA Progress Report 42-87*, vol. July–September 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 92–97, November 15, 1986.
- [16] J. Ruze, "Antenna tolerance theory—A review," *IEEE Proc.*, vol. 54, pp. 663–640, April 1966.

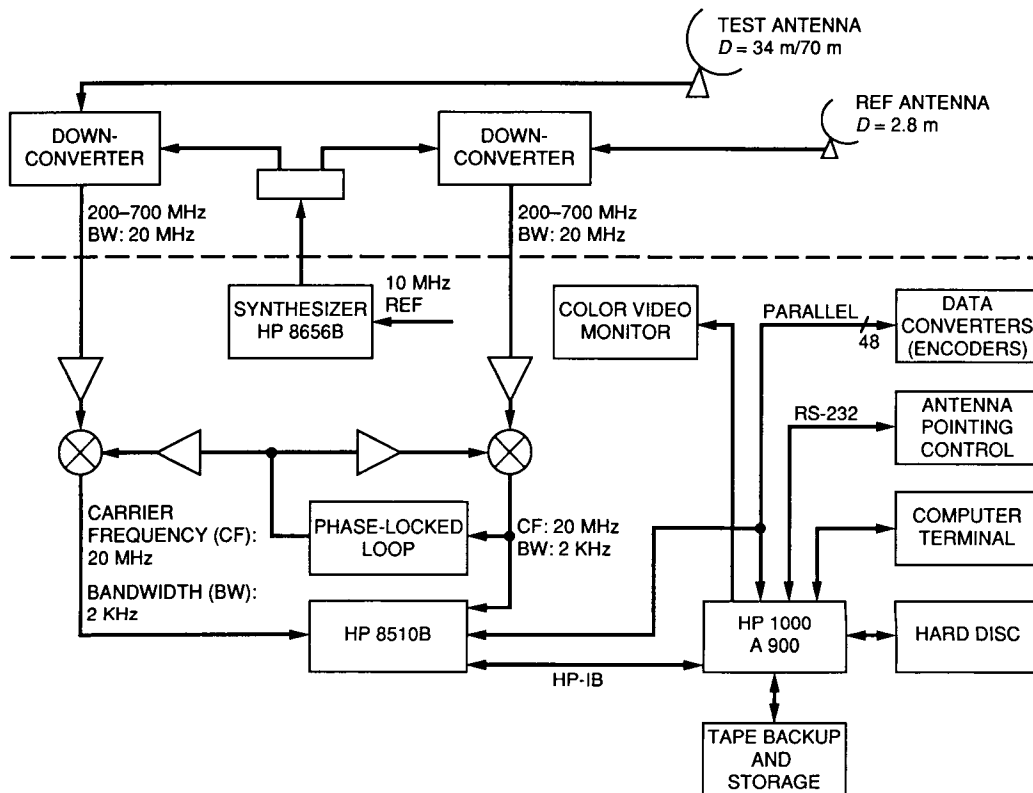


Fig. 1. The holographic measurement using a reference antenna.

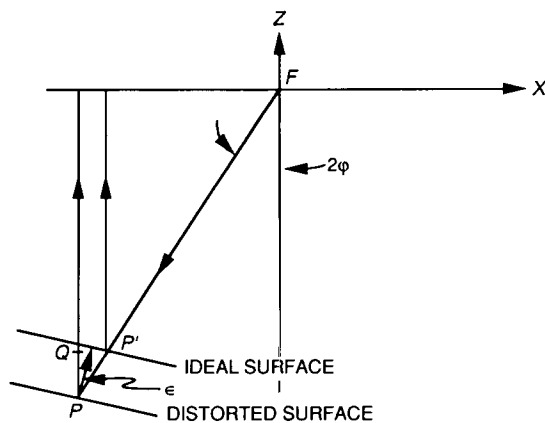


Fig. 2. Surface distortion geometry.

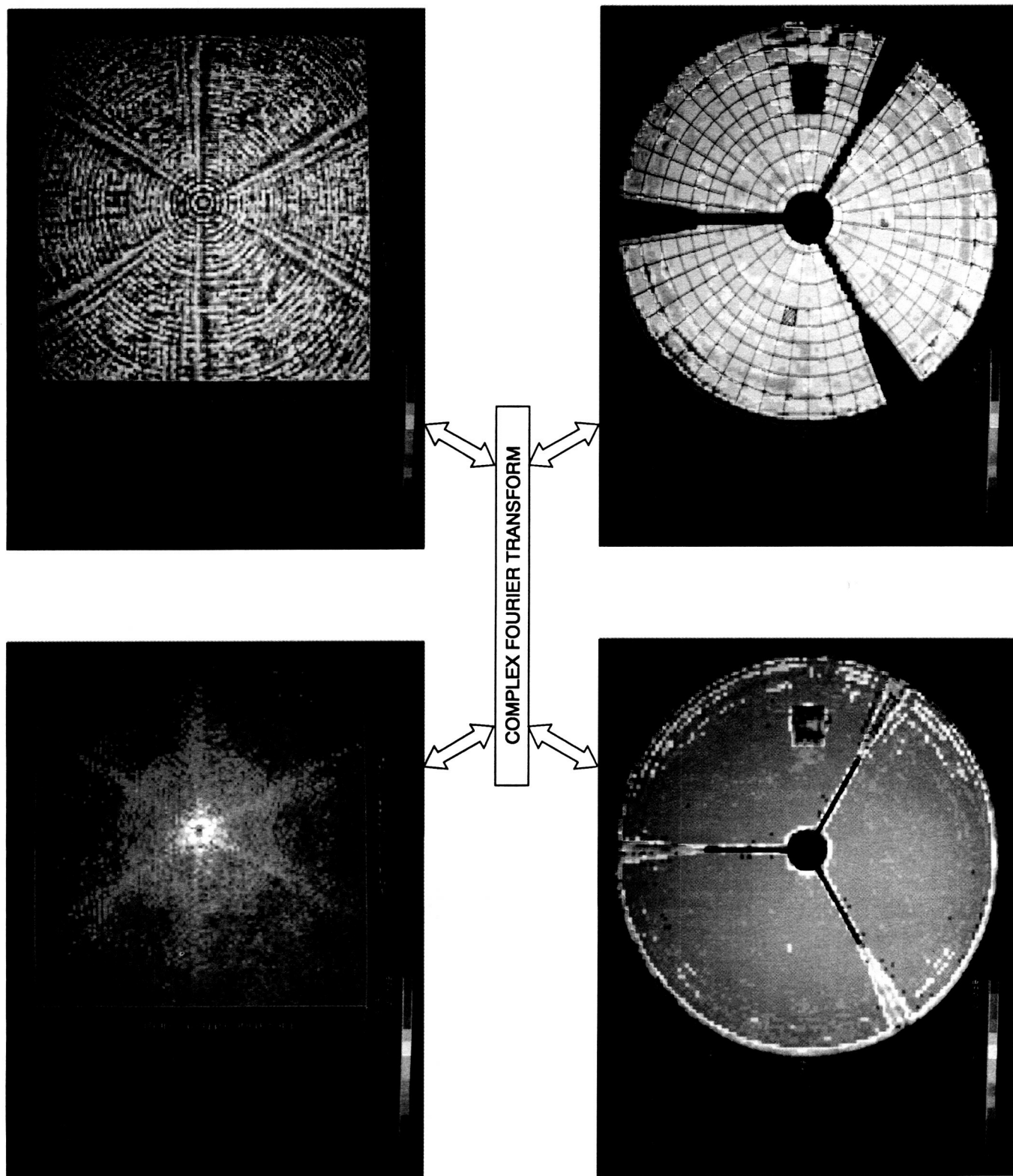


Fig. 3. Complex Fourier transform relating the antenna far-field and aperture functions.

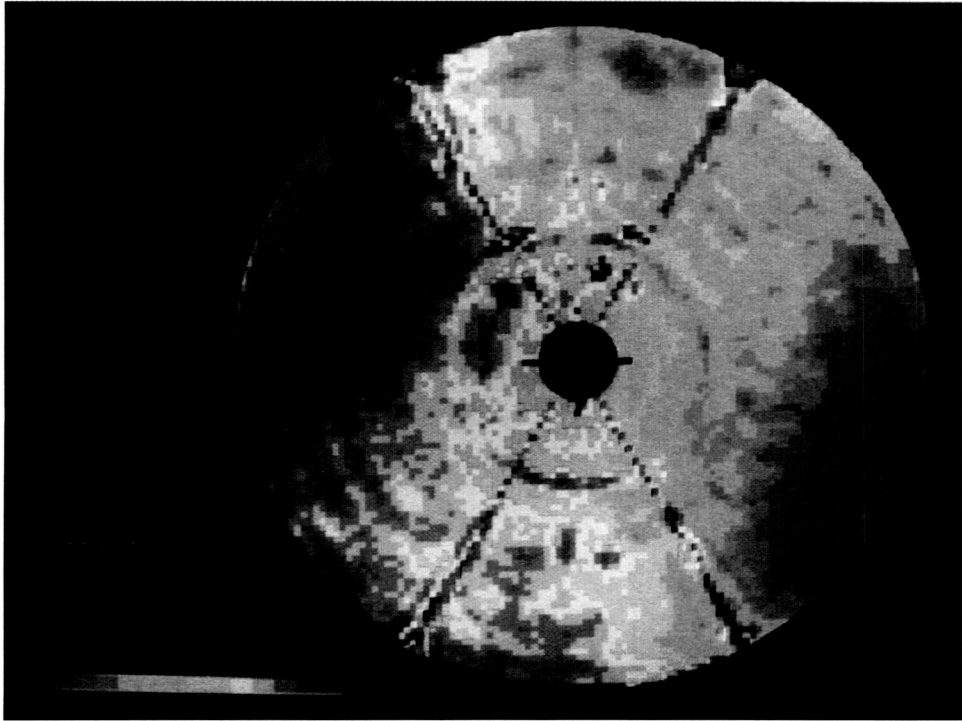


Fig. 4. Phase map before correction for modulo 2π .

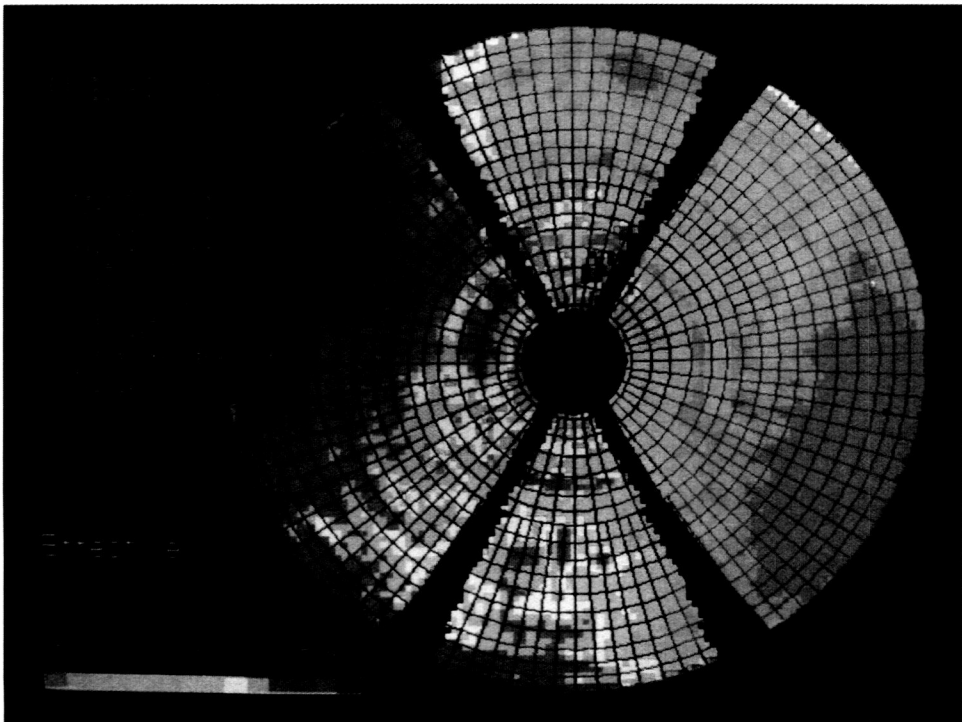


Fig. 5. Phase map after correction for modulo 2π .

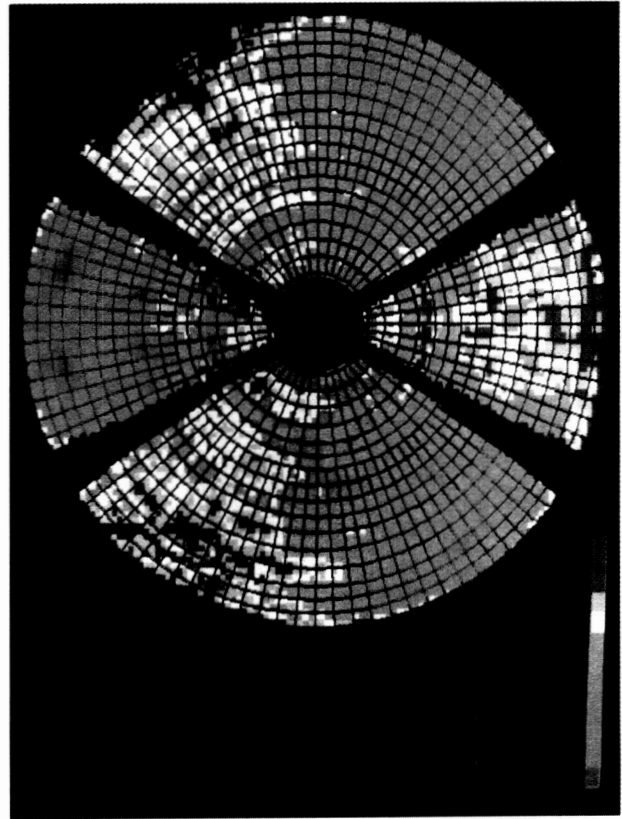
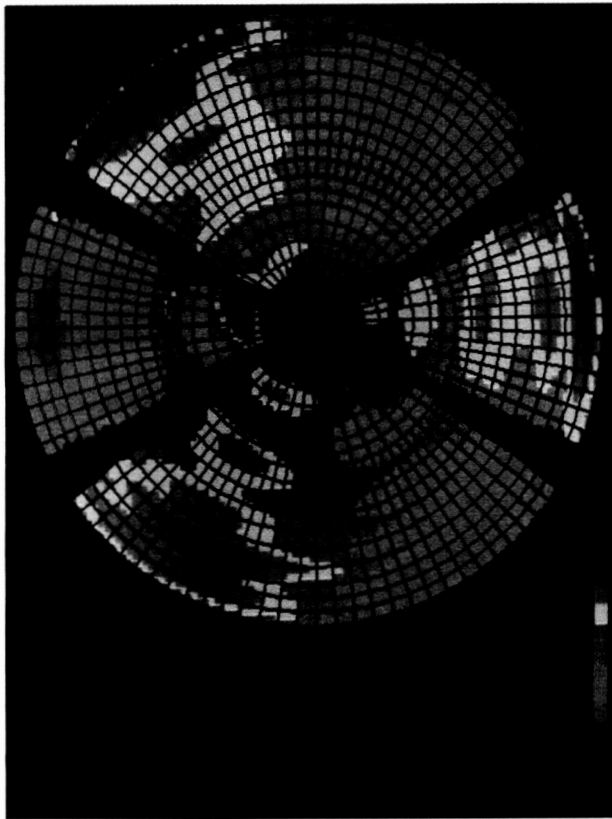
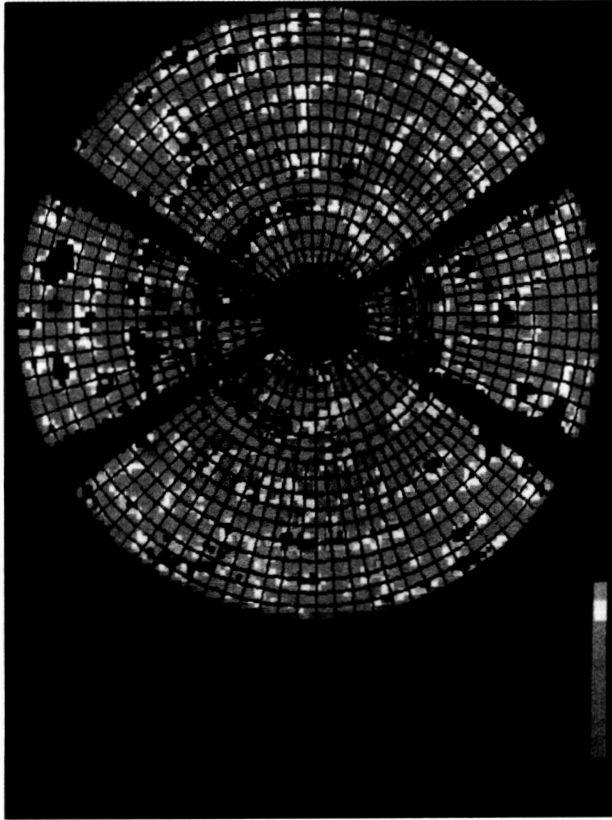


Fig. 6. Modified Jacobi polynomial decomposition separates structural deformation from panel setting errors.

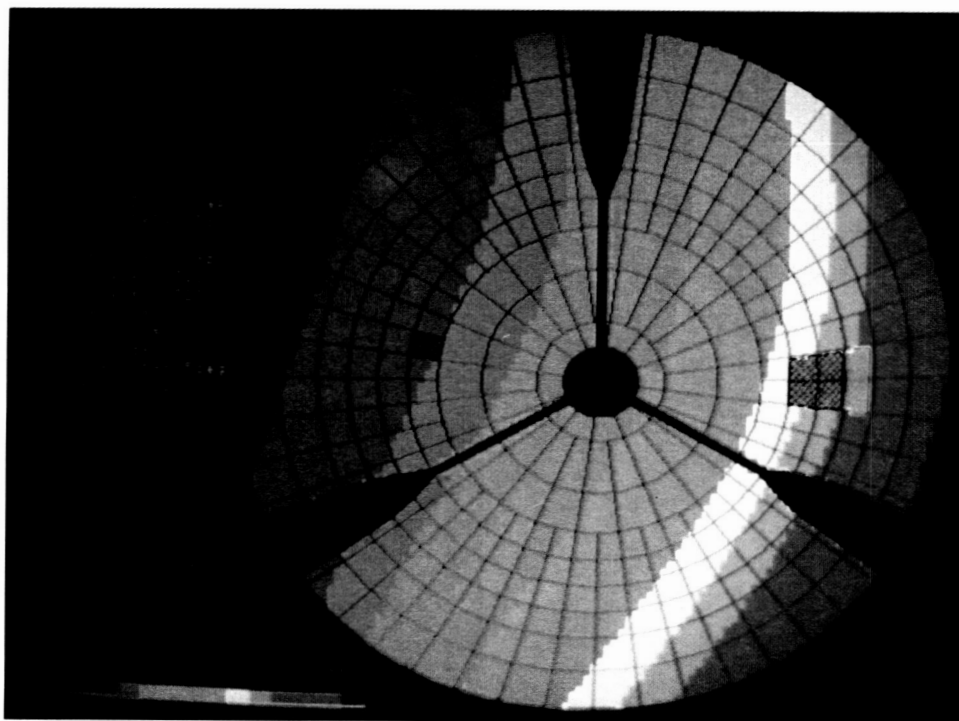


Fig. 7. Phase map due to pointing and subreflector position errors.

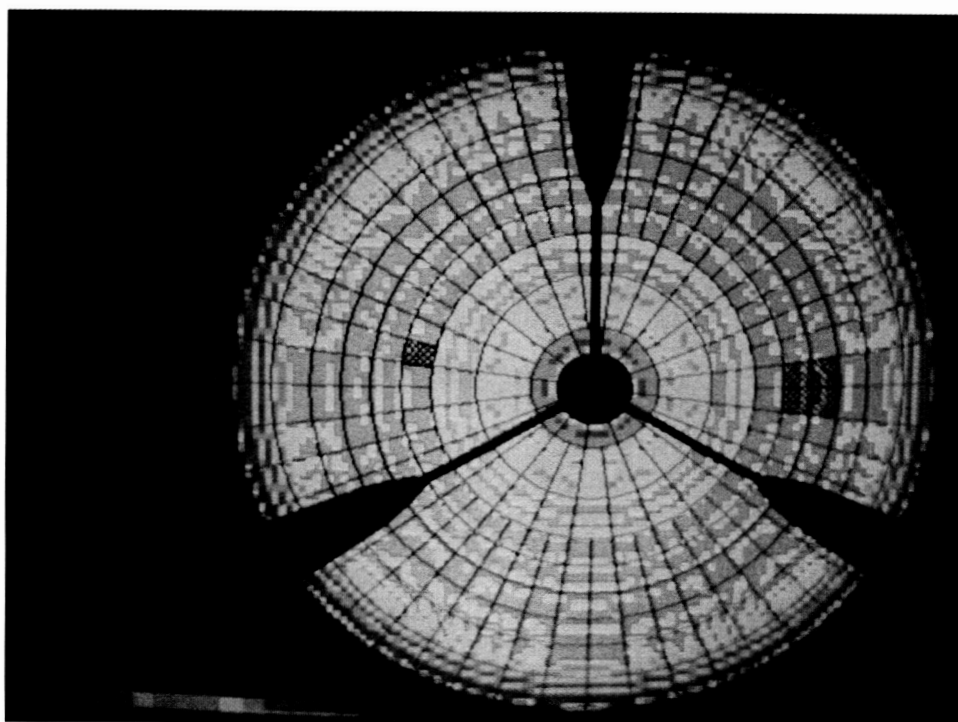


Fig. 8. Feed phase function for DSS-13 34-m antenna at 12.198 GHz.

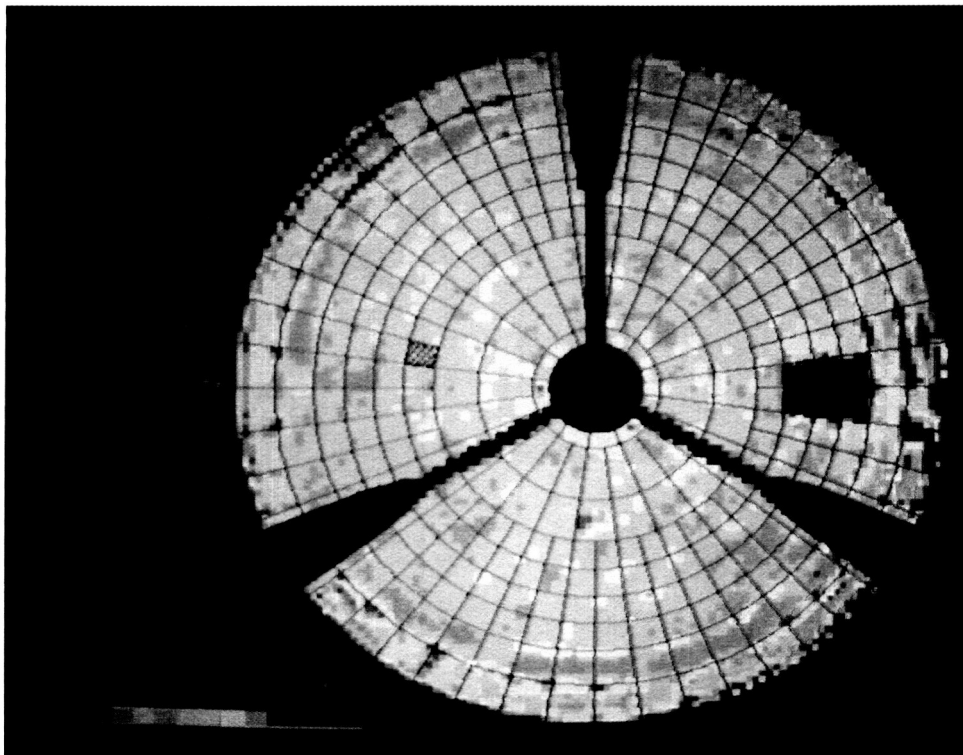


Fig. 9. Effective map.

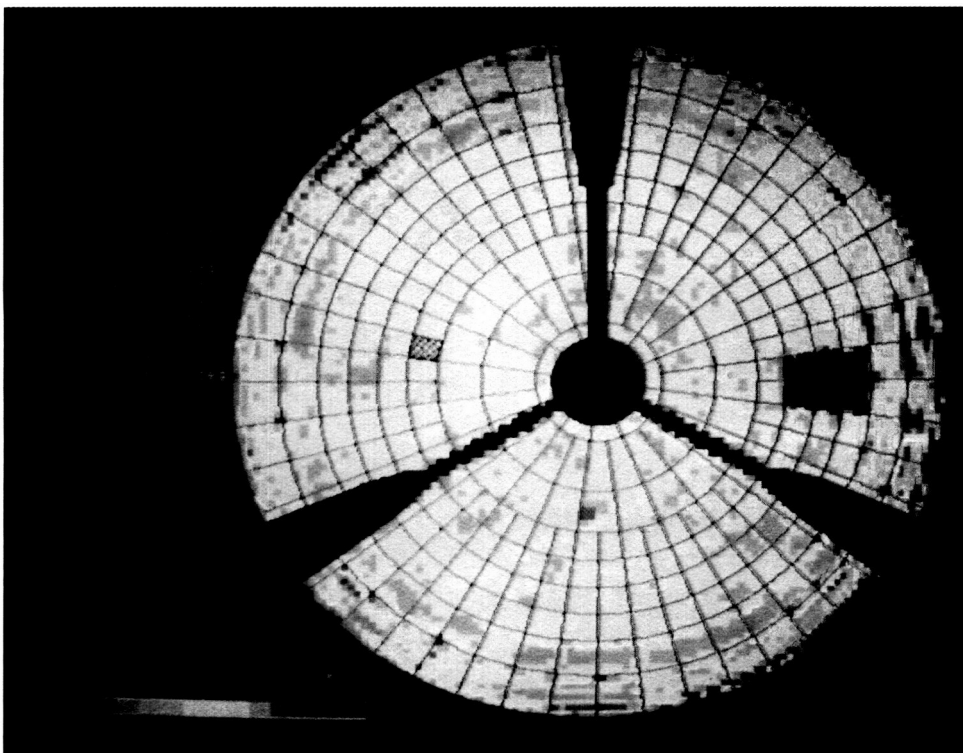


Fig. 10. Mechanical map.

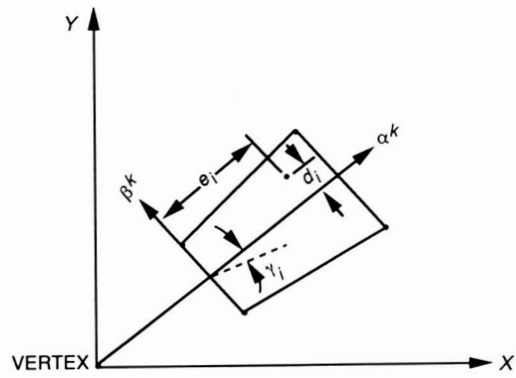


Fig. 11. Rigid-body panel motion modeling.

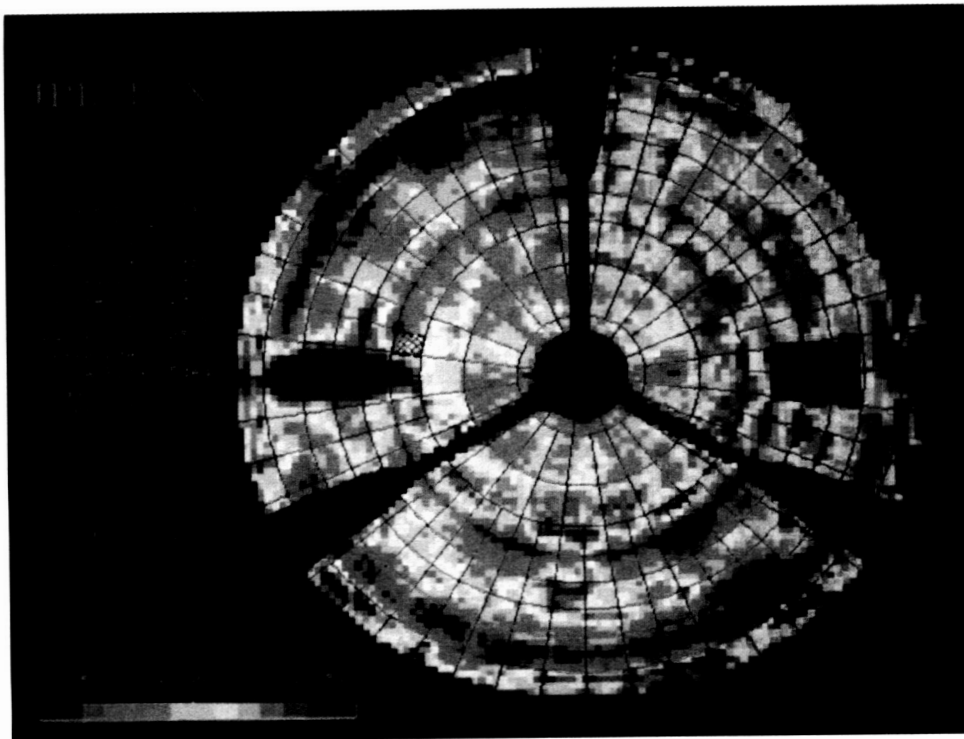


Fig. 12. DSS-13 mechanical surface as found.

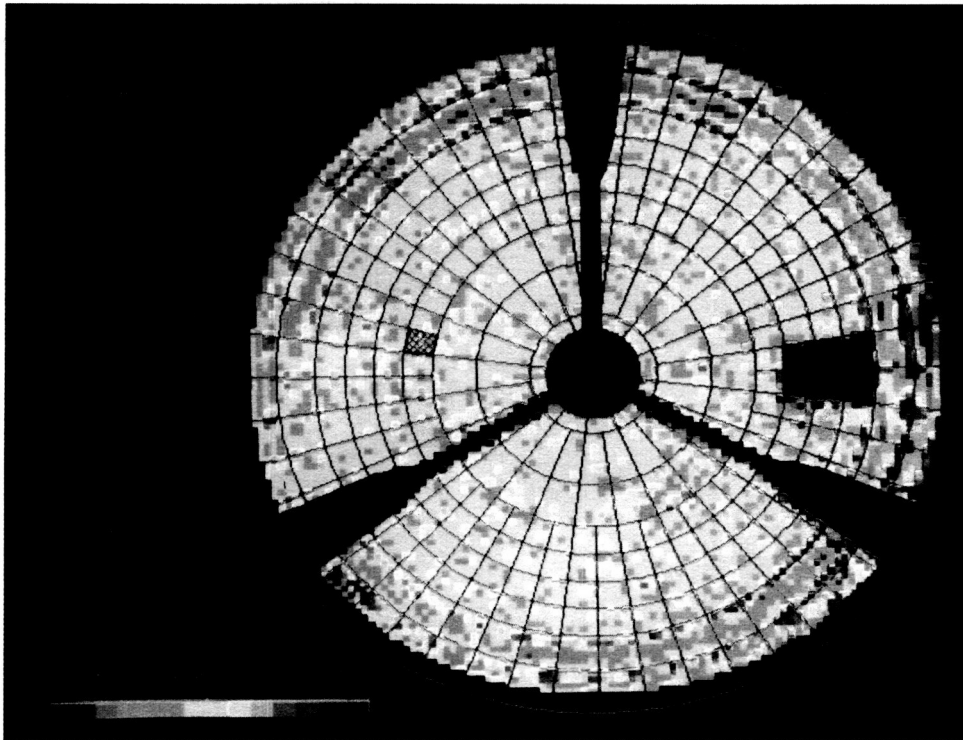


Fig. 13. DSS-13 computer-predicted mechanical surface based on methodology.

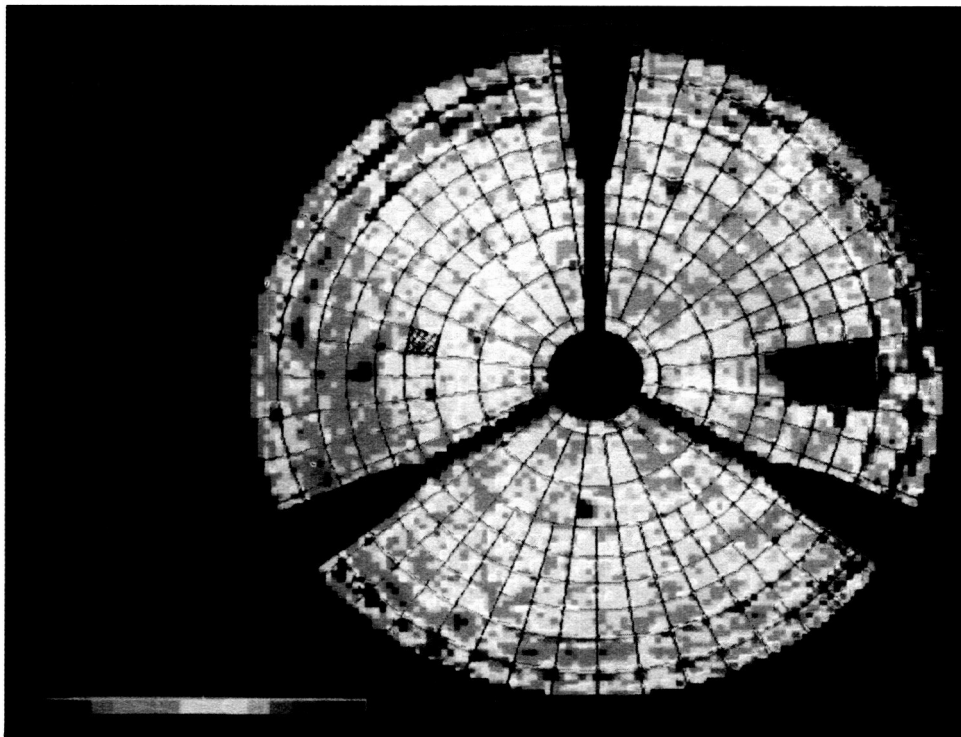


Fig. 14. DSS-13 mechanical surface after panel setting.

SCREW ADJUSTMENTS FOR RING 1						
PANEL	A	B	C	D	E	F
1	0.0464	0.0309	0.0171	0.0065	0.0325	
2	0.0296	0.0047	-0.0006	-0.0047	0.0243	
3	0.0172	0.0071	0.0146	0.0204	0.0248	
4	0.0473	0.0096	0.0084	0.0075	0.0461	
5	0.0296	-0.0044	0.0179	0.0350	0.0521	
6	0.0507	0.0213	0.0110	0.0031	0.0403	
7	0.0593	0.0232	0.0303	0.0358	0.0665	
8	0.0808	0.0837	0.0441	0.0135	0.0407	
9	0.0374	0.0089	0.0229	0.0337	0.0515	
10	0.0654	0.0300	0.0258	0.0226	0.0611	
11	0.0306	0.0230	0.0261	0.0284	0.0337	
12	0.1012	0.1022	0.0494	0.0087	0.0478	

SCREW ADJUSTMENTS FOR RING 2						
PANEL	A	B	C	D	E	F
1	0.0412	0.0224	0.0077	0.0142	0.0275	0.0446
2	0.0150	0.0242	0.0313	0.0020	0.0009	-0.0004
3	-0.0050	0.0027	0.0086	0.0038	-0.0012	-0.0075
4	0.0009	0.0012	0.0015	0.0099	0.0079	0.0054
5	0.0114	0.0116	0.0117	0.0138	0.0132	0.0125
6	0.0188	0.0071	-0.0020	0.0087	0.0156	0.0244

Fig. 15. Panel-setting screw-adjustments (inches) computer listing for DSS-13 34-m antenna derived from scan JPL106 127 × 127.

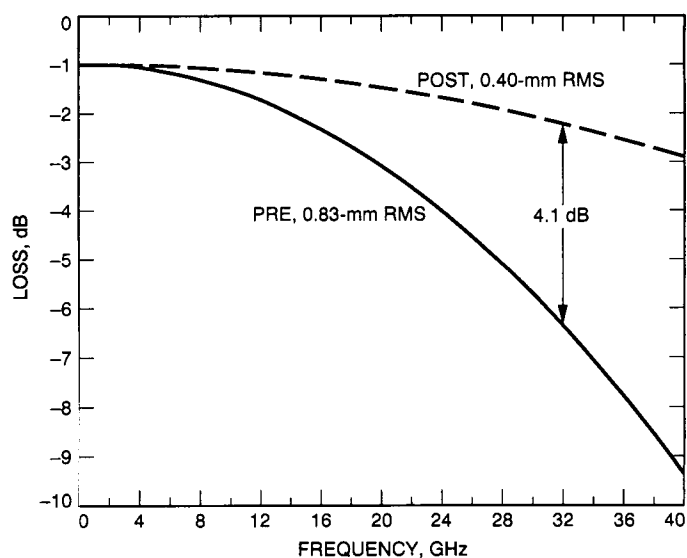


Fig. 16. Estimated DSS-13 surface loss.

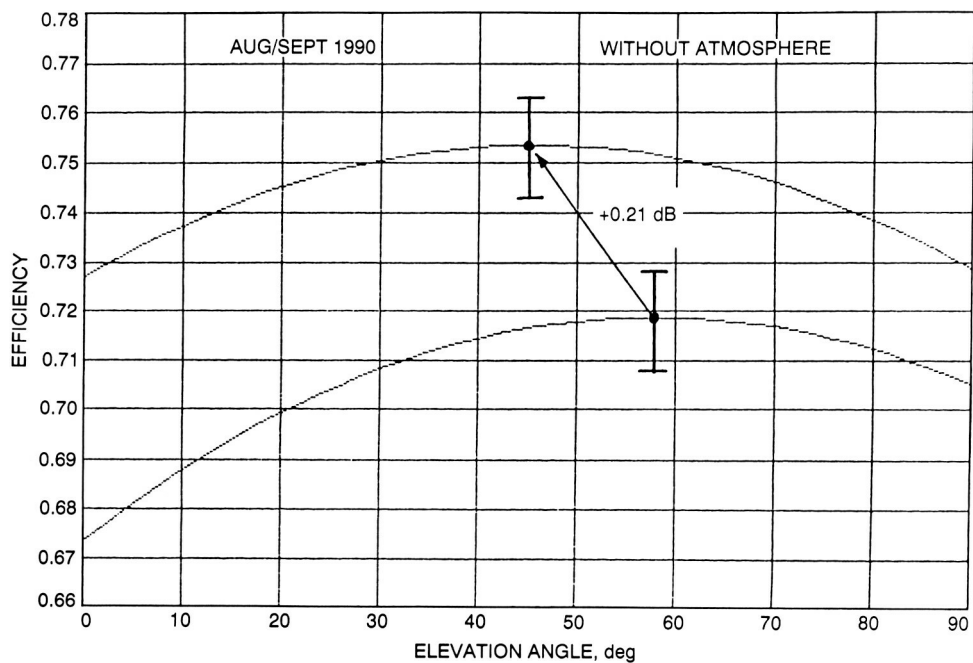


Fig. 17. DSS-13 X-band (8.45-GHz) efficiency at the Cassegrain focus before (bottom) and after (top) application of the holography methodology.

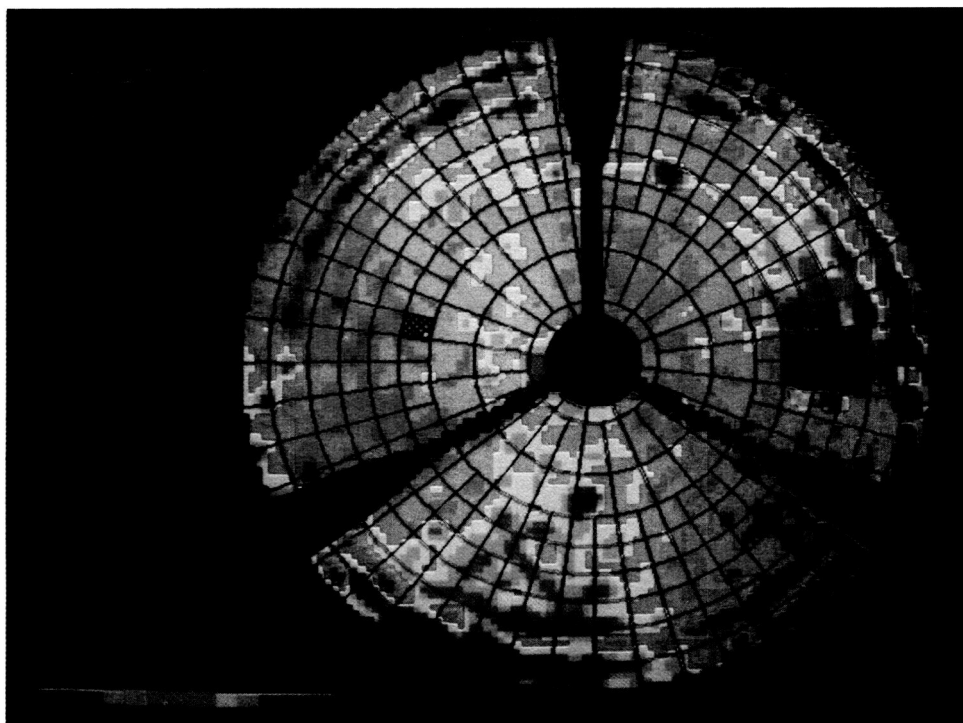


Fig. 18. Gravity deformation signature derived for DSS 13 from measurements at two elevation angles (12.7–46.5 deg).

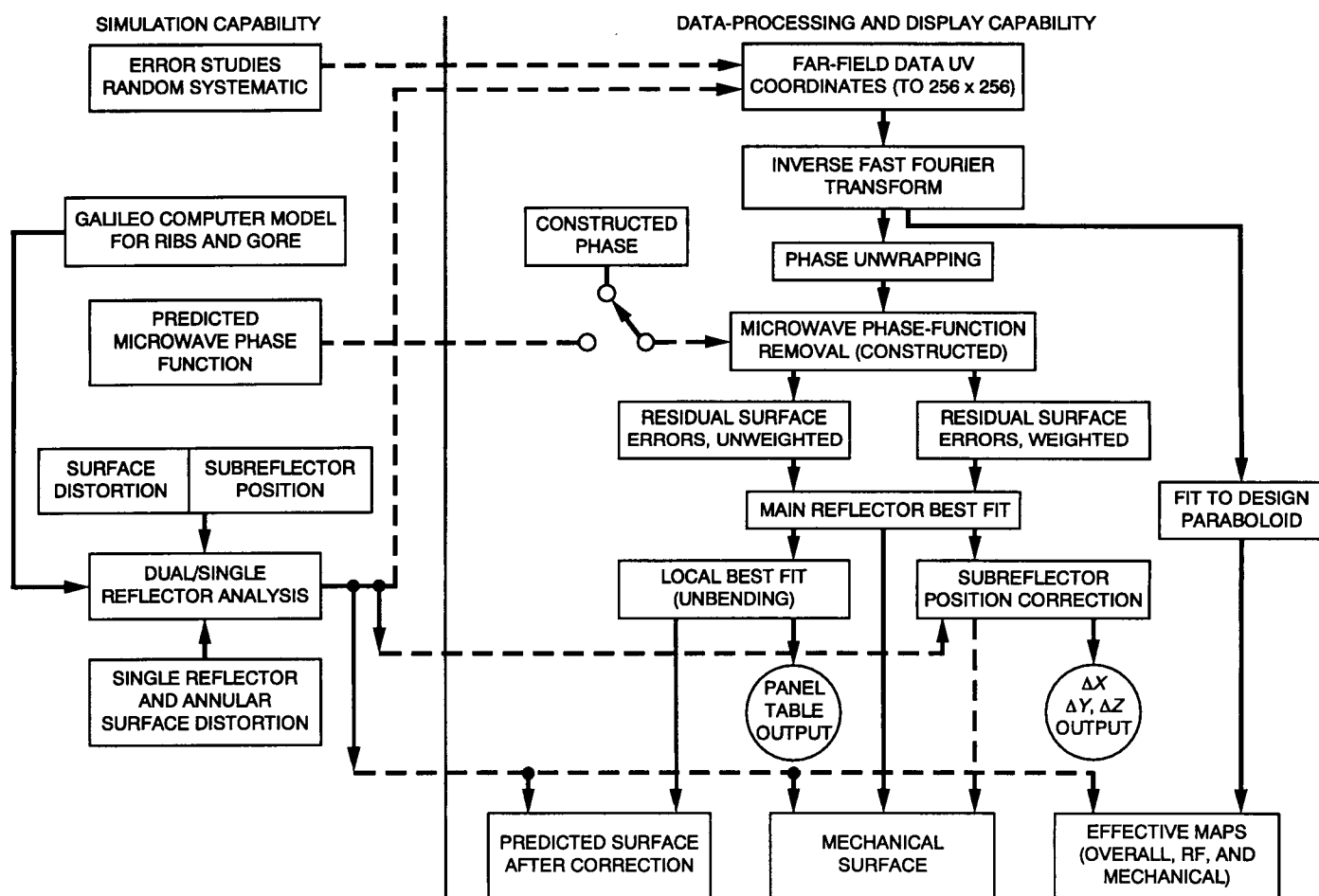


Fig. 19. Simplified block diagram of the microwave holography software methodology.

577-35
77877
P 18
N 9 2 - 2 4 3 2 0
JF 574450

Performance Improvement of DSS-13 34-Meter Beam-Waveguide Antenna Using the JPL Microwave Holography Methodology

D. J. Rochblatt and B. L. Seidel
Ground Antennas and Facilities Engineering Section

This article describes the application of the microwave holography technique to DSS 13. A companion article in this issue (Rochblatt) provides a detailed explanation of the methodology. The project goal of obtaining a rigging-angle surface rms error of 0.43 mm or better was met. The JPL-developed holography algorithms enabled a reduction of the surface error of the DSS-13 antenna from the optically set 0.83-mm axial rms error down to 0.40-mm rms, providing an additional 4.1 dB of performance at 32 GHz.

I. Introduction

The microwave holography technique, when applied to reflector antennas, provides a methodology for examining several critical parameters of the antenna under test [1,2]. The holographically derived information is then used to improve the antenna's RF performance. At DSS 13, more than 4 dB of improvement was obtained at 32 GHz by adjusting the surface of the antenna based on the holographic measurements. Data to aid in the understanding of gravity loading, bypass structural impact, and panel manufacturing contour were obtained. The goal of obtaining a rigging angle surface rms of 0.43 mm or better was achieved. Microwave holography has been demonstrated to be a required tool for initiating, and likely for maintaining, an operational DSN Ka-band (32-GHz) ground antenna capability [3]. Microwave holography has now been applied to all of the 34-m high efficiency (HEF) and 70-m antennas

in the Deep Space Network. The raw data (the observable) for this technique constitute the complex far-field pattern of the antenna under test.

The DSS-13 holography plan called for measurements to be first made at the Cassegrain (f1) focus of the antenna to optimize the subreflector position (Fig. 1), evaluate the antenna main reflector, and, if necessary, reset the surface. Following the f1 measurements, the RF test package [4] was to be moved to the beam-waveguide (f3) focus of the antenna (Fig. 2). The f3 measurements, when differenced from the f1 measurements, would provide insight into the performance of the beam waveguide. Measurements made at several elevation angles would aid in the study of structural deformation due to gravitational effects.

This article describes the measurements made at the f1 focus using a data acquisition system loaned to JPL by the

ORIGINAL CONTAINS
COLOR ILLUSTRATIONS

contractor, Eikontech, Ltd. of England. These data were then transferred to the JPL HP-1000/A900 computer and reduced using JPL-developed (Rochblatt) algorithms. The data analysis software was developed in the Antenna and Microwave Development Group of the Ground Antennas and Facilities Engineering Section [1]. A contractual difficulty precluded making holography measurements at f3.

The information obtained from the JPL analysis algorithms was successfully applied. The subreflector position was verified and the rms surface error of the optically set DSS-13 antenna was substantially reduced. The holographically set surface is better than the specified project goal set forth in the *Functional Requirements Document*¹. The improved antenna surface was verified by both additional holographic and X-band (8.45-GHz) and Ka-band (32-GHz) efficiency measurements. The adjusted surface provides an estimated 4.1-dB improvement in performance at Ka-band (32 GHz), relative to the as-found (optically set) surface. It is estimated that an additional performance improvement of 0.6 dB is possible at 32 GHz over all elevation angles by replacing the panels in rings 8 and 9 with the correctly shaped panels (see the Appendix). It is concluded that holography played a vital and time-efficient role in obtaining a 34-m antenna with an aperture efficiency of 52 percent at Ka-band.

II. Holographic Measurements and Results

In August and September 1990, a total of 24 holography measurements were attempted from the f1 focus of the new DSS-13 beam-waveguide (BWG) antenna. Strong continuous wave (CW) signals obtained from geostationary satellites were used as far-field sources [3,5]. Three different geostationary satellites were successfully scanned, producing four successful high- and medium-resolution data sets at elevation angles of 46.5 deg, 37 deg, and 12.7 deg. The measurements obtained indeed provided the necessary subreflector position information and panel setting information. The measurements also provided a look at the adjusted surface of the antenna, information regarding the exact shape of panels, and information about the gravity performance of the structure at a low elevation angle that revealed the effect of the bypass BWG shroud interaction with the structure. The holographic antenna measurements used satellite signal and ephemeris information supplied by several commercial companies: GTE-operated GSTAR W103, GE-operated SatComm K1, and ComSat-operated Intelsat V.

The inability of the initially implemented servo system to provide precision pointing at the level of the winds experienced during the observations (24 to 48 km/hr, 15 to 30 mph) resulted in the failure of most of the scans. The results of the four successful high- and medium-resolution scans are reported here. Scan JPL106 was a high-resolution scan that provided the first high-resolution look at the surface of the optically set antenna, as well as provided subreflector position error. This scan was used for deriving the required panel setting information and for predicting the best achievable surface rms error. Scan JPL110 was also a high-resolution scan and provided the after-adjustment look at the surface of the antenna. Scans JPL113 and JPL123 provided surface maps of the antenna at two additional elevation angles (37 deg and 12.7 deg, respectively). The last scan, JPL123 at 12.7-deg elevation, provided valuable information for the structural modeling specialists. These four scans are summarized in Table 1.

It is known that the indicated rms error of the antenna surface is affected by the weakly illuminated outer portion of the dish. When the outer edge of the antenna is included in the analysis, the calculated surface rms error is larger than that obtained from analysis of the strongly illuminated portion of the dish. Functionally, the outer 0.6 m of the antenna radius is a noise shield. The rms error obtained from analysis of the central 32 meters of the antenna is therefore more representative of the actual surface than the rms obtained from examination of the full 34-m dish. The rms values for both the full and the central 32 meters of the antenna are included in Table 1. It is estimated that the 1σ error of the rms (central 32 m) is approximately ± 0.05 mm.

The holographic measurement program at DSS 13 started with measurements taken at an elevation angle of about 46.5 deg. High-resolution scan JPL106 supplied the data required for verifying the subreflector position, analyzing the antenna surface, and providing the panel setting information. The surface images, derived from the aperture plane phase, represent the antenna surface deviations from the ideal in the surface normal direction. In the images, the subreflector, the tripod and its shadows, and the bypass beam waveguide are intentionally masked out. The remaining surface is overlaid with an outline of each reflecting panel. The surface error information is shown in pseudocolor with red and blue indicating the high and low deviations, respectively. Figure 3 shows the surface error map of the central 32 meters of the DSS-13 antenna surface as found on August 28, 1990, by scan JPL106. The main reflector surface normal rms error was found to be 0.88 mm (0.77-mm axial) at a resolution of 0.32 m.

¹ R. Sniffin and G. Wood, *DSS-13 Functional Requirements Document*, JPL D-8449 (internal document), Jet Propulsion Laboratory, Pasadena, California, May 13, 1989.

The indicated rms increases as the lateral resolution of the measurement increases. This is an expected result, as there is less area averaging occurring as the resolution increases. The asymptotic or infinite-resolution rms can be estimated by analyzing the scan data at varying resolutions.

Slobin² has estimated that the rms error derived by holography high-resolution (0.32-m) scans is 8 percent lower than the infinite resolution rms. That estimate will be used here for consistency. [Since scan JPL123 (12.7-deg elevation) was taken at medium (0.80 m) resolution, a 19-percent correction is applied to this scan.³]

After accounting for blockage, the holography data analysis software computes both the normal and the axial rms surface errors. Note that the ratio of the axial to normal rms error depends on the precise distribution of the errors. Therefore, the ratio and, hence, the effective slope are somewhat different for each scan. The holography-based analysis computes an effective average slope for the DSS-13 antenna surface of approximately 30 deg. For high-resolution scans, the infinite resolution axial (1/2 path length) rms error can then be approximated by the relationship

$$\text{axial rms error} = \text{normal rms error} \times 1.08 \cos(30 \text{ deg})$$

The above provides a good rule-of-thumb working relationship between the measured high-resolution (0.32-m) normal and the infinite-resolution axial rms surface errors.

Figure 4 shows the predicted surface error map, representing the best achievable surface that would have resulted if the 1716 screws were adjusted precisely as specified by the software. The surface normal rms of this predicted antenna is 0.36 mm at a resolution of 0.32 m. The image reveals that panels in rings 8 and 9 are deformed. A detailed analysis showed that the deformations detected in rings 8 and 9 are due to the fact that the panels used on DSS 13 are of the DSS-15 design shape⁴ (see the Appendix for more details). The decision to use DSS-15 panel molds on DSS 13 was made early during the project for budget

reasons. Note that the present software moves panels as rigid bodies and that further improvement is possible by properly shaping the individual panels.

The panel setting information derived from scan JPL-106 was applied to the surface panel adjusting screws. As a scheduling expedient it was decided to adjust the surface panels by turning the adjusting screws to the nearest 1/8 of a turn (0.16 mm).⁵ Screws requiring adjustment of less than $\pm 1/8$ of a turn were not touched.

The surface error map shown in Fig. 5 was measured on September 7, 1990, after panel resetting (scan JPL110). The deformed panels in rings 8 and 9, as well as many of the other predicted features, are clearly confirmed. The rms surface error achieved by holography-based panel adjustment is 0.45-mm surface normal (0.39-mm axial) at a resolution of 0.32 m. This is equivalent to an infinite-resolution axial rms error of 0.42 mm. This rms error not only exceeds the specification, it also exceeds the goal set forth in the *Functional Requirements Document*. A higher than normal noise level in scan JPL110 leads one to believe that the achieved rms surface error is actually somewhat lower than indicated by this measurement. Indeed, scan JPL113 (Fig. 6), taken at a 37-deg elevation angle, reveals a surface-normal rms error of 0.43 mm and an axial error of 0.37 mm, giving an infinite-resolution axial error of 0.40 mm. Efficiency measurements indicate the 45-deg rigging angle surface to be better than the 37-deg elevation angle surface. One can therefore conclude that, at the rigging angle, the holographically set antenna surface actually has an infinite-resolution axial rms error of slightly less than 0.40 mm.

As shown in Fig. 7, the postholography surface provides a performance improvement of about 0.2 dB at 8.45 GHz, increasing to 4.1 dB at 32 GHz. The DSS-13 surface specifications, along with the corresponding efficiency calculations and measurements, are summarized in Table 2.

Tables 3 through 6 contain X-band and Ka-band calculations of antenna efficiency. The results are obtained by applying Ruze to the holographically obtained axial rms surface error^{6,7} values along with other known block-

² S. Slobin, private communication, Jet Propulsion Laboratory, Pasadena, California.

³ D. J. Rochblatt, *Predicted Performance for DSS-13 at 12.7 Degrees Elevation with Bypass Removal*, JPL IOM 3328-0110 (internal document), Jet Propulsion Laboratory, Pasadena, California, August 28, 1991.

⁴ S. Stewart, D. J. Rochblatt, and B. L. Seidel, *The Effect of Using DSS-15 Main Reflector Panels on the DSS-13 BWG Antenna*, JPL IOM 3328-91-0147 (internal document), Jet Propulsion Laboratory, Pasadena, California, November 8, 1991.

⁵ B. L. Seidel and D. J. Rochblatt, *DSS-13 BWG Antenna Project, Phase I Final Report, Chapter 4: Microwave Holography*, JPL D-8451 (internal document), Jet Propulsion Laboratory, Pasadena, California, May 15, 1991.

⁶ The axial rms surface errors are equal to the 1/2 path-length errors in the familiar form of the Ruze gain-degradation factor $k_g = \exp -[4\pi\epsilon/\delta]^2$. In this expression, ϵ is the 1/2 path error and δ is the wavelength.

⁷ Stewart, Rochblatt, and Seidel, op. cit.

ages and losses. These calculated efficiencies are compared to the radiometrically obtained measurements of efficiency. The surface error map shown in Fig. 8 was derived from medium-resolution (0.80-m) holography measurements made on September 18, 1990, at an elevation angle of 12.7 deg (scan JPL123). The surface-normal rms error at this low elevation angle and resolution is 0.50 mm. The corresponding axial- and infinite-resolution axial rms errors are 0.43 mm and an estimated 0.55 mm, respectively. The holography measurements indicate that the subreflector is displaced from the optimum position by 5 mm laterally and 0.4 mm axially at the 12.7-deg elevation angle (see Table 1). A physical optics computation reveals 0.25 dB of performance loss at 32 GHz for this amount of subreflector displacement. Adding this 0.25 dB of performance loss to the measured 39.4-percent antenna efficiency results in a computed performance of 41.7 percent at this elevation angle.

III. Future Work

The holographically set antenna surface already exceeds the specification set forth in the *Functional Requirements Document*. However, fine-tuning the main reflector surface would yield at least another 0.25 dB of improved Ka-band (32-GHz) performance. Careful subreflector positioning might provide another 0.25 dB of performance at low elevation angles. When developed and applied, panel unbending techniques would further improve aperture efficiency.

By examining the holography data, it was found that using DSS-15 panel molds for manufacturing DSS-13 panels caused an additional 0.6-dB performance loss over all

elevation angles [7]. Thus, even better results could be obtained in the future given another pass of holography adjustments along with some panel unbending.

It is also necessary to return to DSS 13 to make the f3 measurements that were not obtained in Phase 1. Information obtained from these measurements, when differenced from those obtained at f1, would provide an important holographic diagnosis of the BWG effects on the antenna system.

Finally, a microwave holography measurement system will be provided for aligning and maintaining all of the DSN tracking antennas [8]. Such a system is indispensable to an operational Ka-band ground antenna capability in support of future deep space missions and, in fact, is needed even for an X-band network. Microwave holography is a station-time and cost-effective means for providing this support.

IV. Conclusion

The principal f1 holography goal of obtaining a rigging angle surface rms error of 0.5 mm or better at DSS 13 has been met. The JPL-developed holography algorithms, the RF test package, and the concept (and execution) of f1 holography measurements on a BWG antenna worked extremely well. JPL microwave holography enabled reducing the surface error of the DSS-13 antenna from the optically set as-found 0.83-mm axial rms error down to a very respectable 0.40-mm rms. This holographically improved surface not only exceeds the *Functional Requirements Document* specification (0.61 mm), it also exceeds the 0.43-mm goal. The holographically set antenna surface provides an additional 4.1 dB of performance at 32 GHz.

Acknowledgments

The authors thank Dan Bathker for his many helpful technical discussions and contributions. In addition, thanks are due to Manuel Franco from JPL and Paul Wright and Eric Schoessow from Eikontech, Ltd. for helping with the holographic measurements at DSS 13.

References

- [1] D. J. Rochblatt, "A Methodology for Diagnostics and Performance Improvement for Large Reflector Antennas Using Microwave Holography," *Proceedings of the Antenna Measurement Technique Association*, AMTA, Boulder, Colorado, pp. 1-7-1-12, October 7, 1991.
- [2] D. J. Rochblatt, "A Microwave Holography Methodology for Diagnostics and Performance Improvement for Large Reflector Antennas," *TDA Progress Report 42-108*, vol. October–December 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 235–252, February 15, 1992.
- [3] D. J. Rochblatt, "System Analysis for DSN Microwave Antenna Holography," *TDA Progress Report 42-97*, vol. January–March 1989, Jet Propulsion Laboratory, Pasadena, California, pp. 132–157, May 15, 1989.
- [4] T. Y. Otoshi, S. R. Stewart, and M. M. Franco, "A Portable Ku-band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation," *TDA Progress Report 42-107*, vol. July–September 1991, Jet Propulsion Laboratory, Pasadena, California, pp. 73–80, November 15, 1991.
- [5] D. J. Rochblatt and Y. Rahmat-Samii, "Effects of Measurement Errors on Microwave Antenna Holography," *IEEE Transactions on Antennas and Propagation*, vol. 39, no. 7, pp. 933–942, July 1991.
- [6] D. J. Rochblatt, Y. Rahmat-Samii, and J. H. Mumford, "DSN Microwave Antenna Holography Part II: Data Processing and Display of High Resolution Effective Maps," *TDA Progress Report 42-87*, vol. July–September 1986, Jet Propulsion Laboratory, Pasadena, California, pp. 92–97, November 15, 1986.
- [7] D. J. Rochblatt and B. L. Seidel, "DSN Microwave Antenna Holography," *TDA Progress Report 42-76*, vol. October–December 1983, Jet Propulsion Laboratory, Pasadena, California, pp. 27–42, February 15, 1984.
- [8] D. J. Rochblatt, "System Requirements for NASA Deep Space Network Microwave Antenna Holography," paper presented at the International Symposium on Antenna and Propagation, Tokyo, Japan, August 22–25, 1989.

Table 1. Results of DSS-13 microwave holography at f1.

	Scan JPL106	Scan JPL110	Scan JPL113	Scan JPL123
Satellite	GSTAR W103	GSTAR W103	SatComm K1	Intelsat V
Elevation angle	46.5	46.5	37.0	12.7
Measurement frequency, GHz	12.198	12.198	12.198	11.701
Surface	Mechanical	Mechanical	Mechanical	Eff ^a
Array size	127 x 127	127 x 127	127 x 127	51 x 51
Resolution, m	0.32	0.32	0.32	0.80
RMS, mm				
Surface normal				
Full antenna	1.07	0.58	0.71	0.66
Central 32 m	0.88	0.45	0.43	0.50
Axial				
Central 32 m	0.77	0.39	0.37	0.43
Infinite resolution	0.83	0.42	0.40	0.55
Estimated surface error				
Loss, dB				
Ruze				
2.30 GHz	0.028	0.007	0.006	0.012
8.34 GHz	0.375	0.096	0.087	0.160
32.00 GHz	5.375	1.376	1.248	2.360
Estimated subreflector				
Position error, mm				
X	-0.53	+0.28	+1.07	-1.88
Y	-1.06	-1.41	-3.07	-4.72
Z	+1.10	+1.13	+0.66	+0.39

^a Eff: Effective surface error map without removal of the feed/subreflector phase function [6].

Table 2. DSS-13 rigging angle surface specifications, efficiency calculations and measurements at f1.

Antenna surface	RMS, mm	Calculated/Measured aperture efficiency, percent	
		8.45 GHz	32.0 GHz
As found	0.83	71.7/71.9	21.6/ ^a
Specified	0.61	74.6/ ^a	38.3/ ^a
Goal	0.43	76.4/ ^a	53.6/ ^a
After reset	0.40	76.6/75.4	56.0/52.3
Potential (existing panels)	0.36	76.9/ ^a	59.2/ ^a
Potential (replacing panels in rings 8 and 9)	0.28	77.4/ ^a	65.4/ ^a

Notes:

- (1) Efficiency estimates include 0.07-dB I^2R loss at 8.45 GHz and 0.27-dB I^2R loss at 32 GHz.
- (2) Both the 8.45-GHz and the 32-GHz efficiency estimates include 186.6 m² of aperture blockage.
- (3) The measured efficiency values (supplied by Slobin) are derived from radiometric measurements.
- (4) Given a perfect main reflector surface, the estimated X- and Ka-band efficiencies would be 78.2 percent and 74.7 percent, respectively. This difference is due to the different I^2R losses at the two frequencies.

^a N/A or no measured data available.

Table 3. Comparison of holography-based efficiency estimate with measured value: X-band postadjustment at rigging angle (using the 37-deg infinite resolution rms value).

DSS-13 34-m BWG antenna, 8.45-GHz at 46-deg elevation		
Gross area		907.9 m ²
Less		
Noise shield	64.7 m ²	
S/R blockage	9.0 m ²	
Strut blockage	74.4 m ²	
Bypass blockage	23.5 m ²	
Illumination taper	15.0 m ²	
Subtotal	186.6 m ²	
Ruze loss, 0.40 mm	14.3 m ²	
S/R position loss	Negligible	
Subtotal effective area		707.0 m ²
I ² R loss, X-band, 0.07 dB	11.3 m ²	
Net effective area		695.7 m ²
Area efficiency		
Estimated		76.6 percent
Measured		75.4 percent

Table 5. Comparison of holography-based efficiency estimate with measured value: Ka-band postadjustment at rigging angle (using the 37-deg infinite resolution rms value).

DSS-13 34-m BWG antenna, 32-GHz at rigging angle		
Gross area		907.9 m ²
Less		
Noise shield	64.7 m ²	
S/R blockage	9.0 m ²	
Strut blockage	74.4 m ²	
Bypass blockage	23.5 m ²	
Illumination taper	15.0 m ²	
Subtotal	186.6 m ²	
Ruze loss, 0.40 mm	180.2 m ²	
S/R position loss	Negligible	
Subtotal effective area		541.1 m ²
I ² R loss, X-band, 0.27 dB	32.6 m ²	
Net effective area		508.5 m ²
Area efficiency		
Estimated		56.0 percent
Measured		52.3 percent

Table 4. Comparison of holography-based efficiency estimate with measured value: X-band postadjustment at 12.7 deg.

DSS-13 34-m BWG antenna, 8.45-GHz at 12.7-deg elevation		
Gross area		907.9 m ²
Less		
Noise shield	64.7 m ²	
S/R blockage	9.0 m ²	
Strut blockage	74.4 m ²	
Bypass blockage	23.5 m ²	
Illumination taper	15.0 m ²	
Subtotal	186.6 m ²	
Ruze loss, 0.40 mm	26.8 m ²	
S/R position loss	Negligible	
Subtotal effective area		694.5 m ²
I ² R loss, X-band, 0.07 dB	11.2 m ²	
Net effective area		683.3 m ²
Area efficiency		
Estimated		75.3 percent
Measured		74.0 percent

Table 6. Comparison of holography-based efficiency estimate with measured value: Ka-band postadjustment at 12.7 deg.

DSS-13 34-m BWG antenna, 32-GHz at 12.7-deg elevation		
Gross area		907.9 m ²
Less		
Noise shield	64.7 m ²	
S/R blockage	9.0 m ²	
Strut blockage	74.4 m ²	
Bypass blockage	23.5 m ²	
Illumination taper	15.0 m ²	
Subtotal	186.6 m ²	
Ruze loss, 0.55 mm	302.4 m ²	
S/R position loss, 0.25 dB	25.7 m ²	
Subtotal effective area		393.2 m ²
I ² R loss, X-band, 0.27 dB	26.2 m ²	
Net effective area		367 m ²
Area efficiency		
Estimated		40.4 percent
Measured		39.4 percent

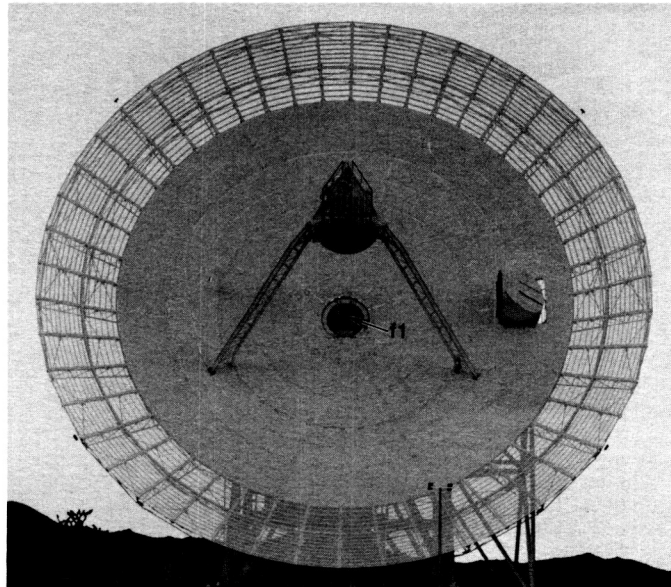


Fig. 1. DSS-13 34-m BWG antenna showing the Ku-band horn feed at the f_1 position at the Cassegrain focus.

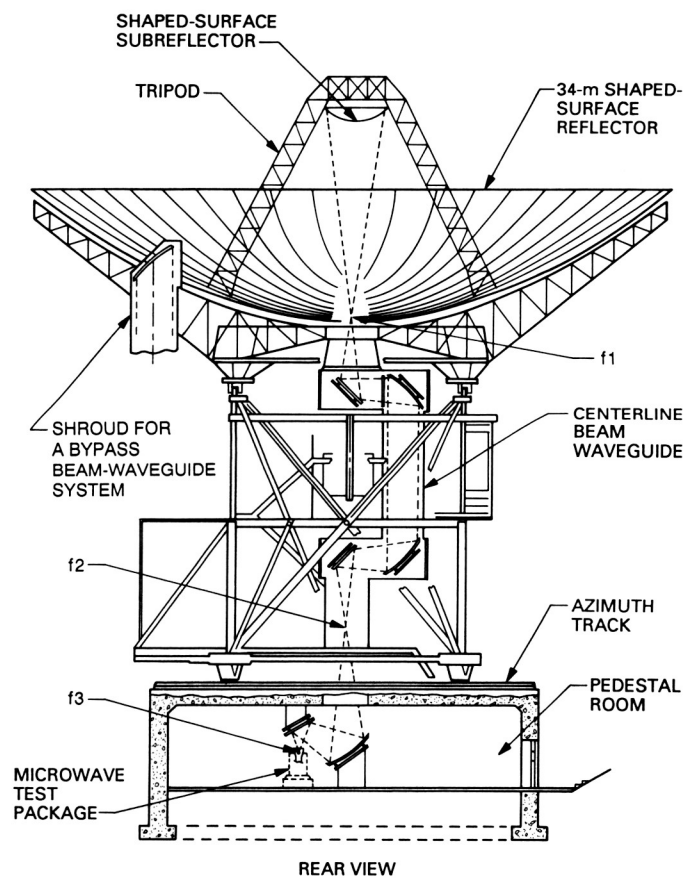


Fig. 2. DSS-13 antenna showing f_3 beam waveguide focus.

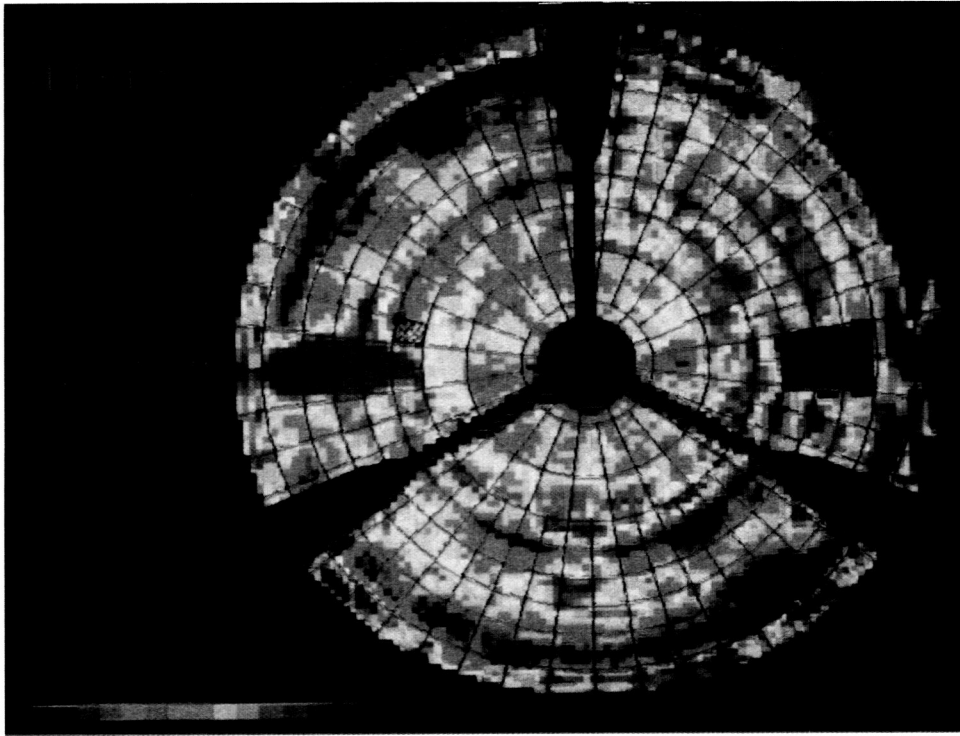


Fig. 3. High-resolution (0.32-m) error map of the central 32 m of the DSS-13 antenna surface at 46-deg elevation, before panel setting, as derived from scan JPL106 (August 28, 1990).

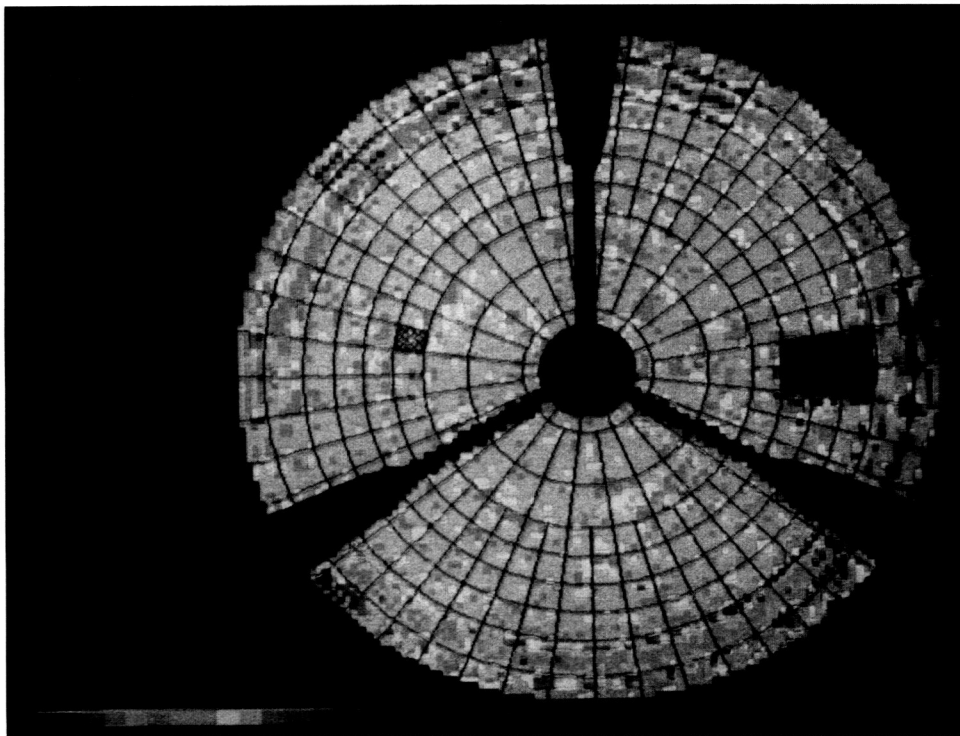


Fig. 4. Predicted surface error map derived from scan JPL106. This represents the best achievable rigging angle surface that would have resulted if the 1716 screws were adjusted precisely as specified by the software.

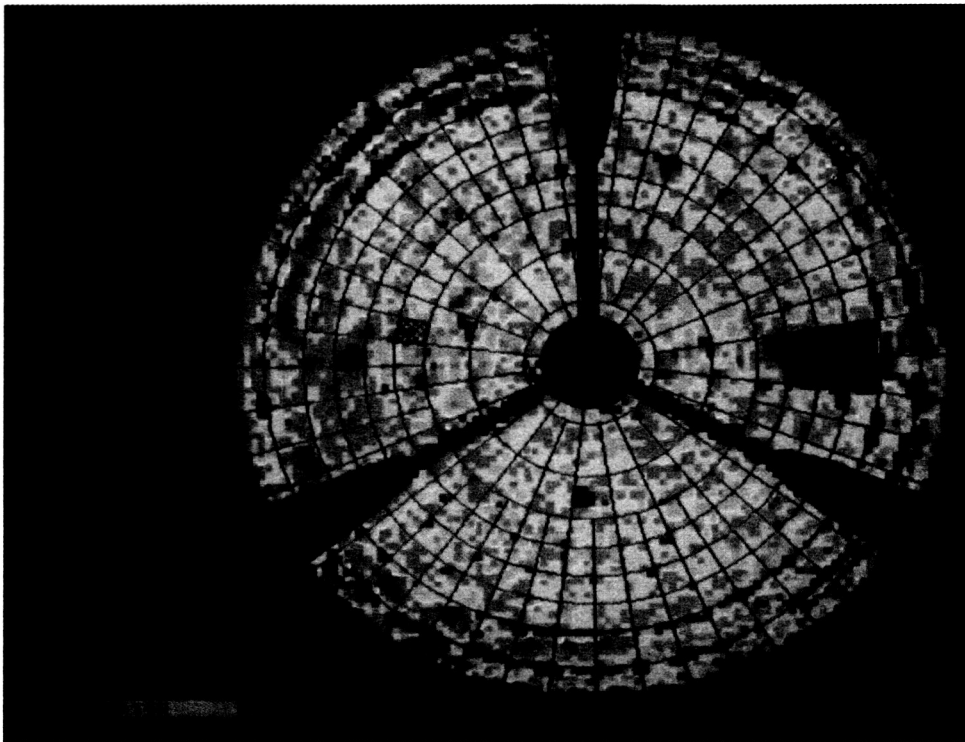


Fig. 5. High-resolution (0.32-m) error map of the central 32 m of the DSS-13 antenna surface at 46-deg elevation, after panel setting, as derived from scan JPL110 (September 7, 1990).

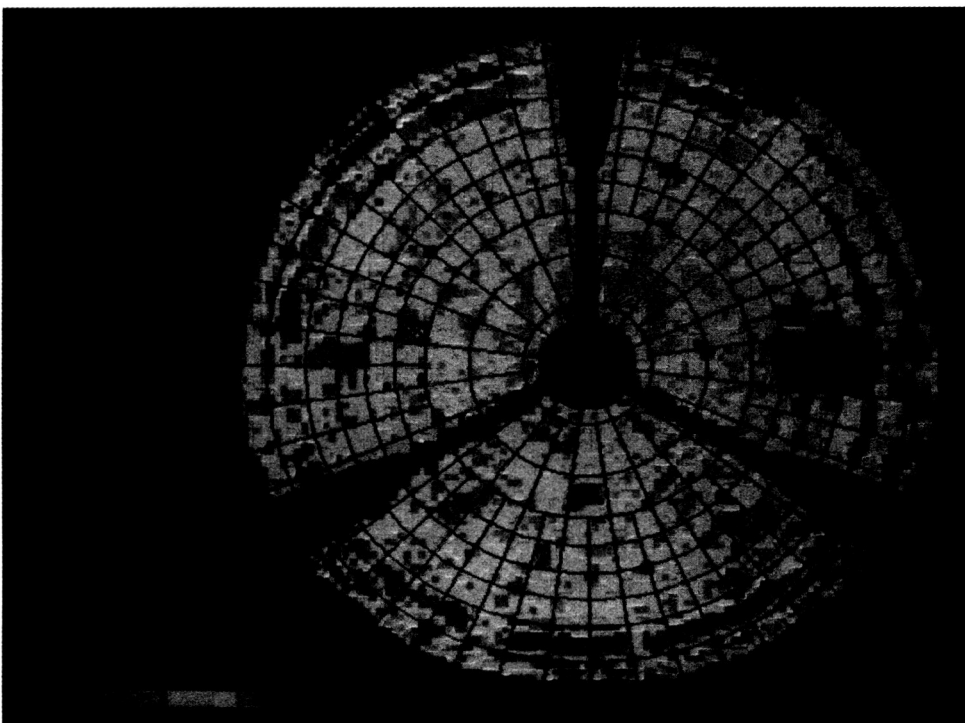


Fig. 6. High-resolution (0.32-m) error map of the central 32 m of the DSS-13 antenna surface at 37-deg elevation, after panel setting, as derived from scan JPL113 (September 11, 1990).

C-4

ORIGINAL PAGE
COLOR PHOTOGRAPH

ORIGINAL PAGE
COLOR PHOTOGRAPH

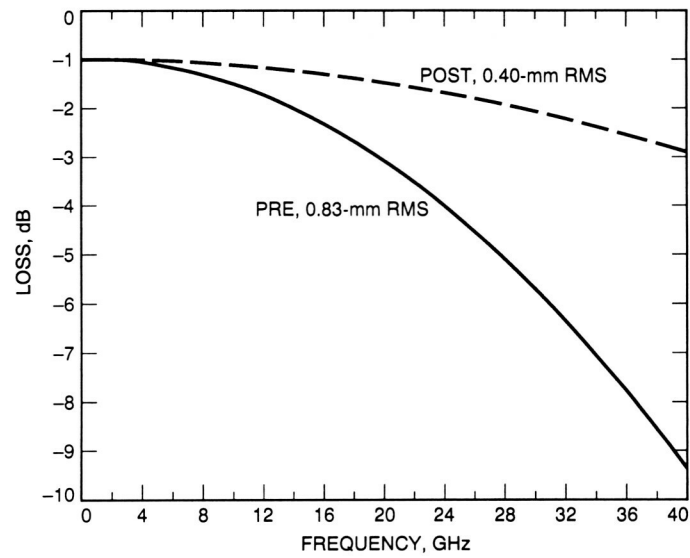


Fig. 7. Loss due to surface error versus frequency. The loss indicated in these curves is based on aperture blockage and surface roughness. No attempt was made to include I^2R losses. An estimated improvement in antenna performance of approximately 4.1 dB was achieved at 32 GHz after panel resetting.

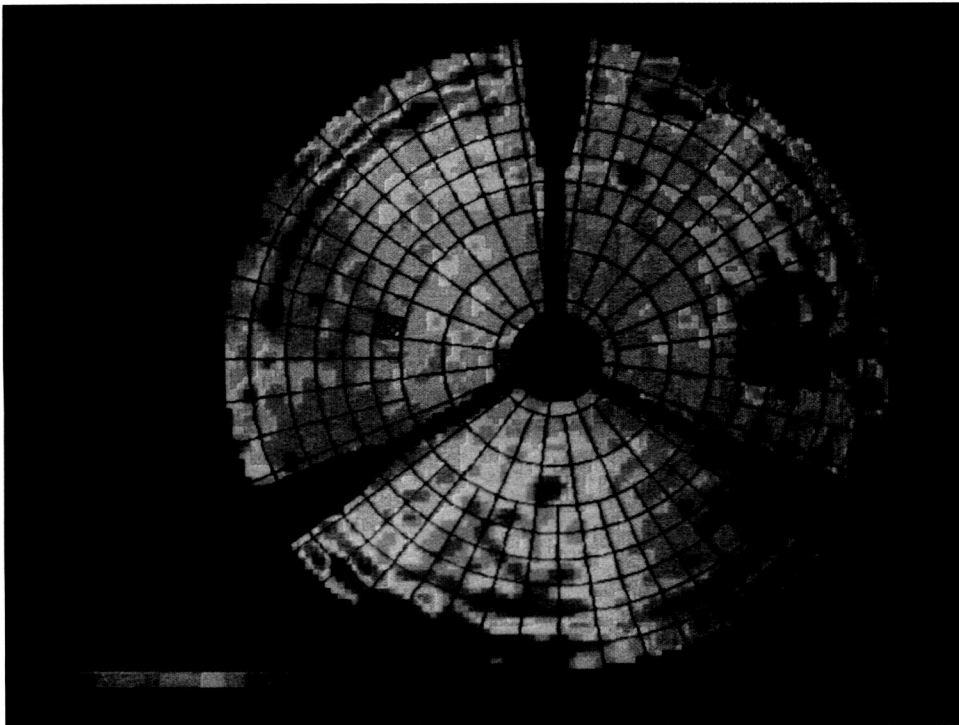


Fig. 8. Medium (0.80-m) resolution error map of the central 32 m of the DSS-13 antenna surface at 12.7-deg elevation, after panel setting, as derived from scan JPL123 (September 18, 1990).

Appendix

The Effect of Using DSS-15 Main Reflector Panel Molds for Fabrication of DSS-13 Panels

S. Stewart, D. J. Rochblatt, and B. L. Seidel

In 1988, when the new DSS-13 BWG antenna project was still in the planning stages, it was decided that the main reflector panels for the DSS-13 antenna should be made from the available molds that were used to make the main reflector panels for the DSS-15 antenna. The differences in the shape of the panels were thought to be minor, and it was believed they would not significantly affect the required (per the *Functional Requirements Document*) performance of the new research and development antenna.

Looking at each of the nine panel rings individually and assuming that the panels on the DSS-13 antenna were made accurately from the DSS-15 manufacturing contours, the panels were mathematically best-fitted to the DSS-13 design contour. The axial errors between these two contours were calculated for each of the nine panel rings by subtracting the reference DSS-13 required shape from the DSS-15 panel contour.

The errors in the first seven panel rings are minor and cause no significant performance loss at 32 GHz. The errors in panels 8 and 9 (Figs. A-1 and A-2, respectively) are much more sizable and contribute noticeably to RF performance degradation at 32 GHz. (It should be noted that the outer half of panel 9 is a noise shield and should not be viewed as contributing to the RF gain performance.)

The next step was to look at the holography data measured at the f1 focal point in order to determine if these mechanical errors of as much as 0.79 mm were detected. Several holography data sets were examined, comprising measurements made both prior to and after panel resetting, as well as computed prediction models. Figure A-3 shows the mechanical surface error map that was obtained from the holographic measurements made on the DSS-13 antenna on August 28, 1990, prior to panel resetting. This map describes the surface errors of the as-found optically set antenna. The magnitude of the various mechanical surface errors across the dish make detection of the errors due only to the misshaped panels impossible from this map. The predicted best achievable antenna surface was derived from this mechanical surface error map by the holography software and is shown in Fig. A-4. The software makes no attempt to untwist distorted panels, but rather adjusts

them, in a least-squares fit, as rigid bodies. The result of the fitting represents the best rms antenna surface achievable without unbending panels. The remaining surface errors are due primarily to the shape of the panels.

Figures A-5 and A-6 present the inner seven rings and the outer two rings of the predicted surface, respectively. The deformed panels in rings 8 and 9 are clearly noticeable and account for the large rms surface error of 0.60 mm for that part of the dish compared to the much lower rms surface error of 0.24 mm for the inner seven rings. It should be noted that 0.24 mm is the rms of the inner seven rings of the best or predicted surface, and can only be obtained when the panel-screw-adjustment computer listings are followed exactly. In actuality, 0.28-mm rms was achieved for this portion of the dish surface. It is interesting to note that the rss difference of the actual (0.28 mm) and the best achievable surfaces (0.24 mm) is approximately equal to the screw adjusting accuracy, hence, 1/8 of a turn (0.16 mm versus 0.14 mm).

Figures A-7 and A-8 are plots for rings 8 and 9, respectively, of the constructed average symmetric phase error function from the predicted surface error map and the analytically derived feed phase function at the measurement frequency of 12.198 GHz. The figures clearly indicate that the signature of the feed phase function, which is frequency dependent, is properly absent in the constructed function derived from the predicted surface error map. Therefore, the confidence is high that the feed phase function has been accurately accounted for. Figures A-9 and A-10 show the holographically derived surface errors for rings 8 and 9, respectively, overlaid on the predicted mechanical surface errors for these panels. These plots show good agreement between the manufacturing contour and the holography measurements.

The current antenna surface has an rms error of 0.28 mm for the inner seven panel rings, 0.60 mm for rings 8 and 9 together, and 0.40 mm for all of the antenna, excluding the noise shield.

The errors in the DSS-13 BWG antenna surface in the outer two panel rings have been measured accurately and are noticeable in 32-GHz performance. Estimates have

been made as to the possible gain improvement that would be obtained if these panels were replaced by panels that had surface errors on the order of those seen on the inner seven panel rings. The outer two rings incorporate 37 percent of the aperture area, not including the noise

shield. If this area of the antenna were to have its rms error decreased from 0.60 mm to 0.28 mm, the increase in antenna gain would be 0.6 dB at 32 GHz. The present Ka-band area efficiency of 52 percent at f1 would increase to 60 percent if the 0.6-dB opportunity is pursued.

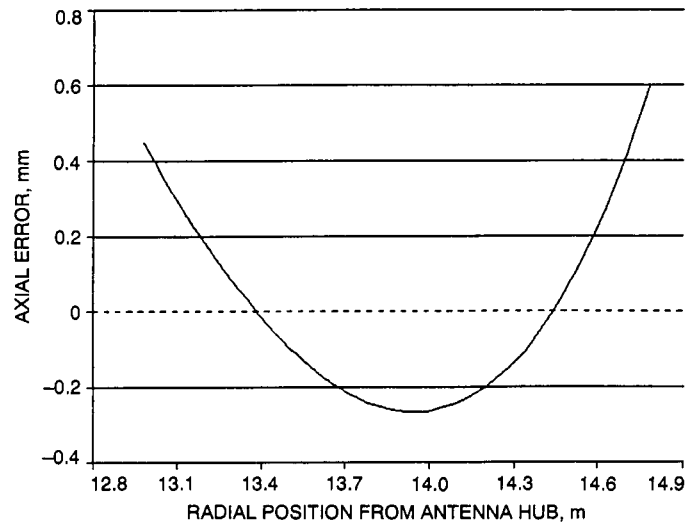


Fig. A-1. Mechanical error in DSS-13 main reflector surface panel 8.

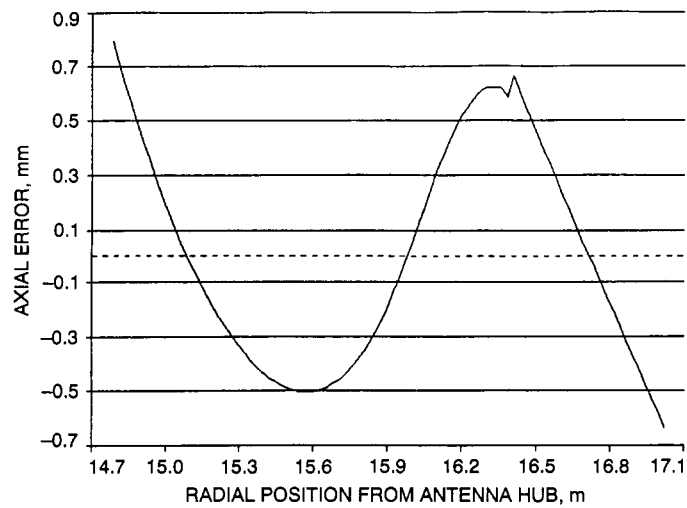


Fig. A-2. Mechanical error in DSS-13 main reflector surface panel 9.

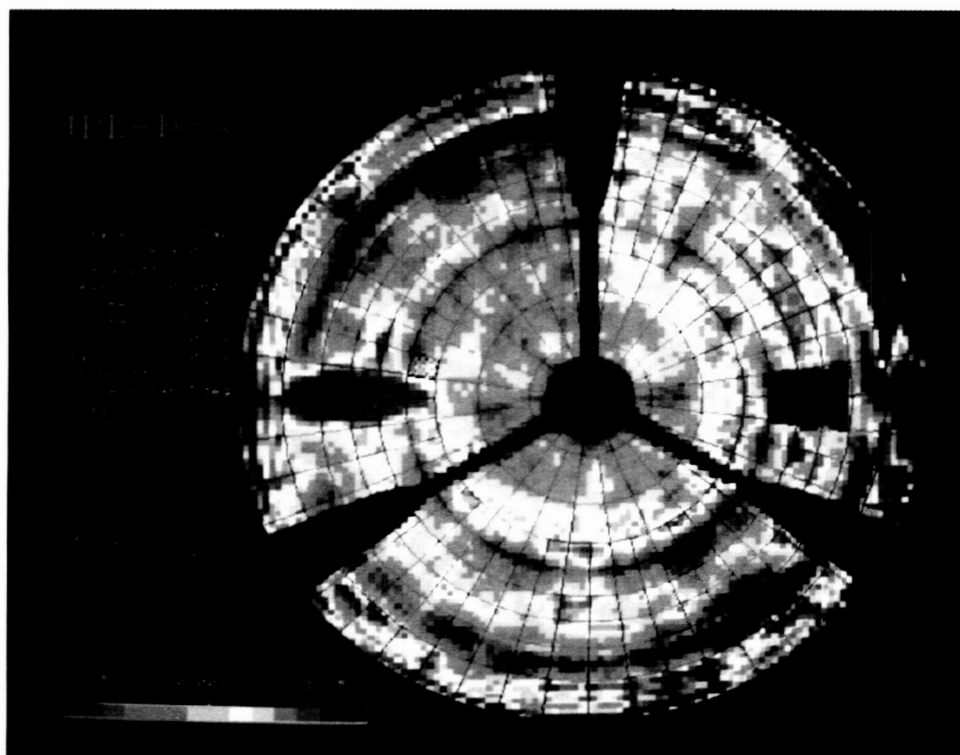


Fig. A-3. High-resolution mechanical error map of DSS-13 antenna surface at 46.5-deg elevation before panel setting.

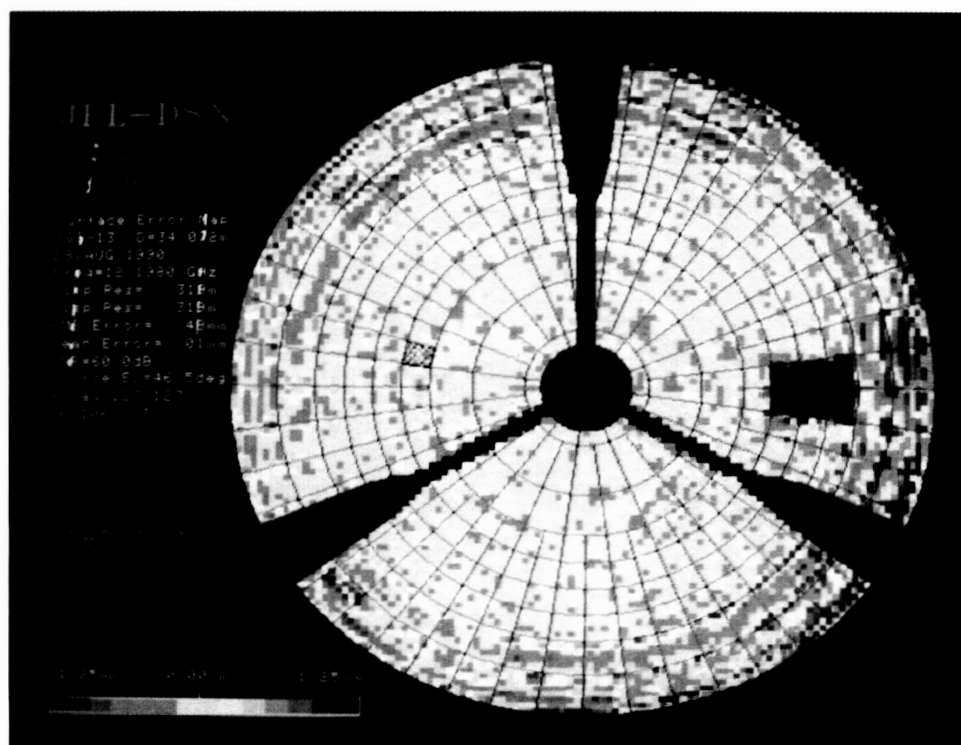


Fig. A-4. Predicted surface error map representing the best achievable rigging angle surface resulting from precision adjustment of set screws.

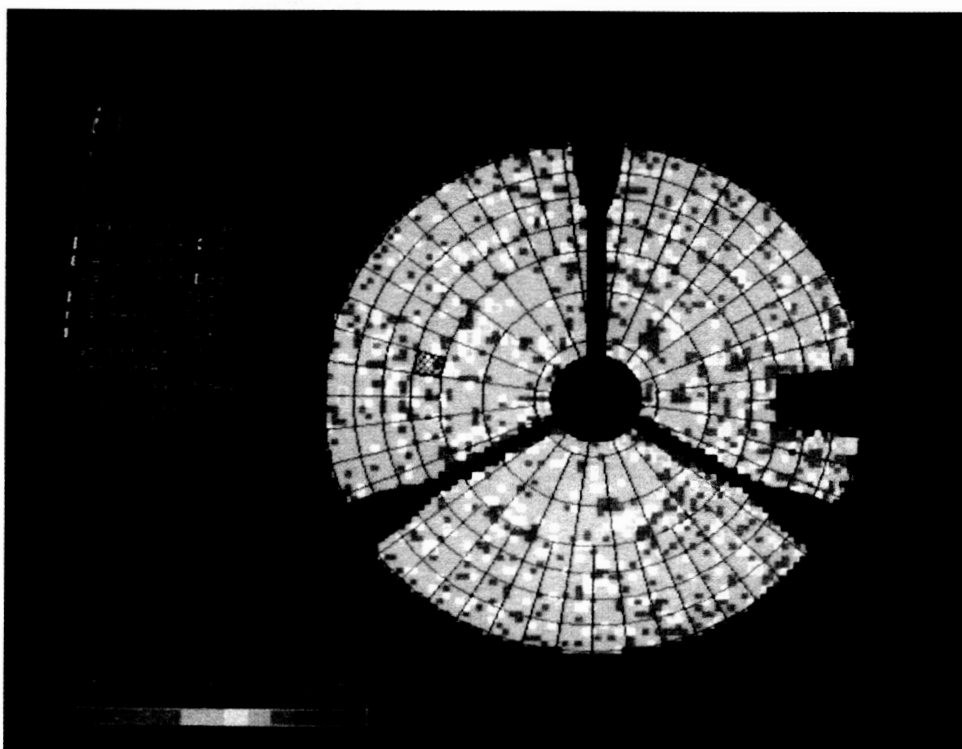


Fig. A-5. Inner seven panel rings with infinite resolution axial rms error of 0.24 mm.

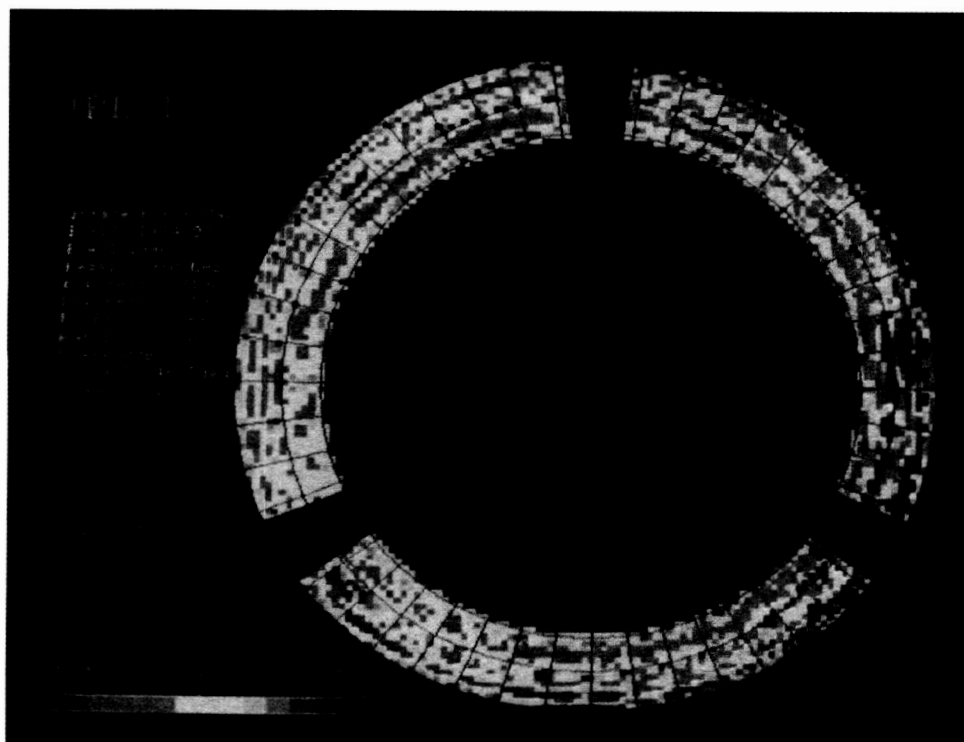


Fig. A-6. Panel rings 8 and 9 with infinite resolution axial rms error of 0.60 mm.

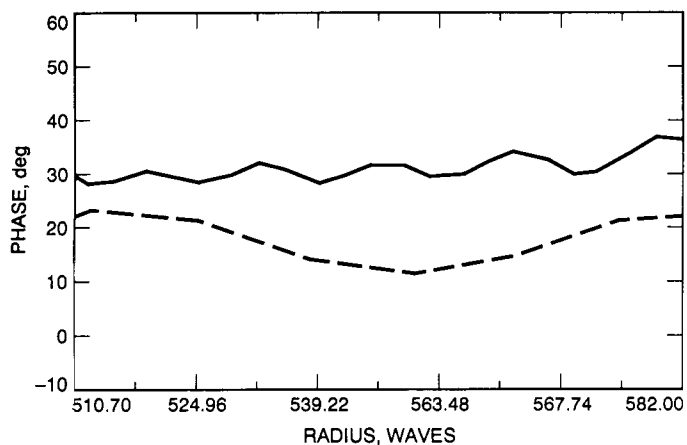


Fig. A-7. Panel ring 8—feed phase function (solid line) and constructed phase function (dashed line).

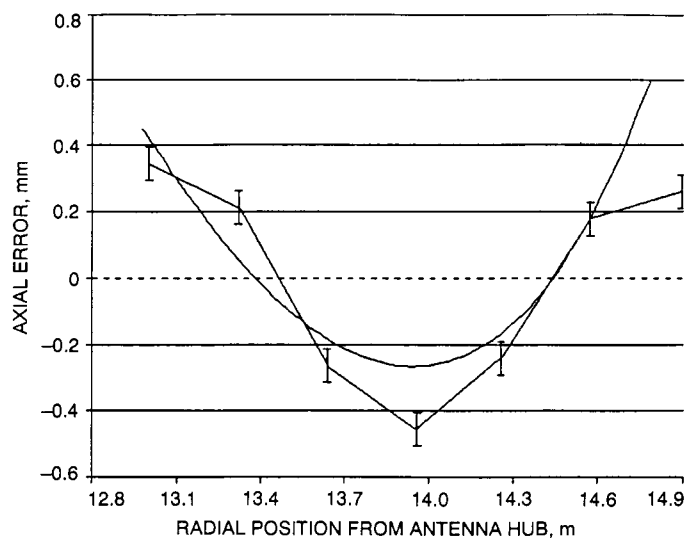


Fig. A-9. Mechanical error in panel ring 8 with holographically derived average phase error function.

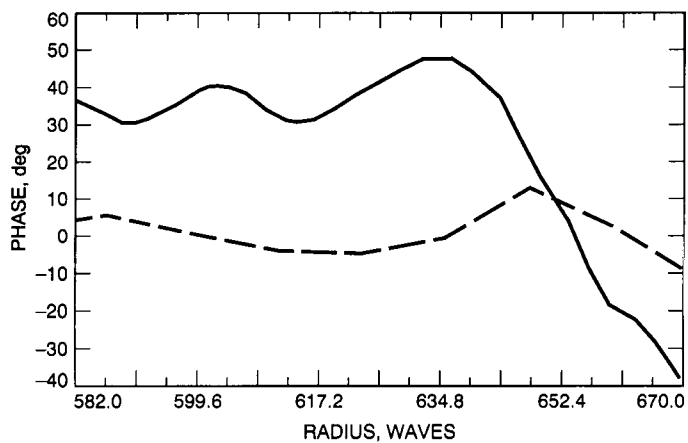


Fig. A-8. Panel ring 9—feed phase function (solid line) and constructed phase function (dashed line).

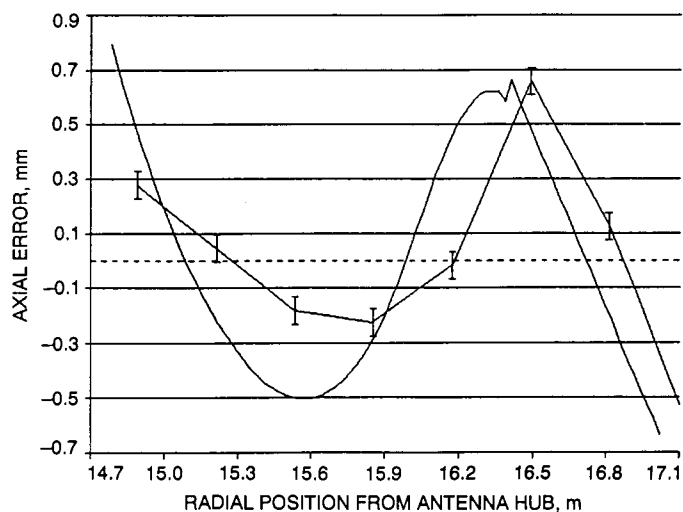


Fig. A-10. Mechanical error in panel ring 9 with holographically derived average phase error function.

N92-24321

J5574450

DSN G/T_{op} and Telecommunications System Performance

C. Stelzried, R. Clauss, and W. Rafferty

TDA Technology Development

S. Petty

Radio Frequency and Microwave Subsystems Section

This article provides an intersystem comparison of present and evolving DSN microwave receiving systems. Comparisons of the receiving systems are based on the widely used G/T_{op} figure of merit, which is defined as antenna gain divided by operating system noise temperature. In 10 years, it is expected that the DSN 32-GHz microwave receiving system will improve the G/T_{op} performance over the current 8.4-GHz system by 8.3 dB. To compare future telecommunications system end-to-end performance, both the receiving systems' G/T_{op} and spacecraft transmit parameters are used. Improving the 32-GHz spacecraft transmitter system is shown to increase the end-to-end telecommunications system performance an additional 3.2 dB, for a net improvement of 11.5 dB. These values are without a planet in the field of view (FOV). A Saturn mission is used for an example calculation to indicate the degradation in performance with a planet in the FOV.

I. Introduction

The ground-based microwave receiving system is an important and continuously evolving element of the DSN spacecraft-to-ground telecommunications system. The recent development of a research and development (R&D) 34-m beam waveguide (BWG)¹ antenna is part of a plan to prepare for DSN operation at Ka-band (the Ka-band deep space-to-ground frequency allocation is 31.8 to 32.3 GHz). This article provides an intersystem comparison based on the receiving system figure of merit, G/T_{op} (defined as antenna gain divided by operating system noise tem-

perature), and the spacecraft transmitter power and antenna area. These parameters are used with an end-to-end telecommunications system performance (TSP) equation, which is defined below. The comparisons represent current understanding of DSN antennas and low noise amplifiers.

The receiving system figure of merit, G/T_{op} , is determined by parameters that can be measured conveniently and separately. Throughout this article, G/T_{op} characterizes the receiving microwave antenna and receiver front-end; that is, cosmic noise, atmospheric effects, antenna, maser, and the follow-up amplifier.

The microwave receiving system figure of merit is [1]

$$M = G/T_{op} \quad (1)$$

¹M. Britcliffe, L. S. Alvarez, D. A. Bathker, P. W. Cramer, T. Y. Otsoshi, D. J. Rochblatt, B. L. Seidel, S. D. Slobin, S. R. Stewart, W. Veruttipong, and G. E. Wood, *DSS-13 Beam Waveguide Antenna Project*, JPL D-8451 (internal document), Jet Propulsion Laboratory, Pasadena, California, May 15, 1991.

where

$M = G/T_{op}$ = receiving system figure of merit

M (dB) = G (dB) - T_{op} (dB)

G = antenna gain (relative to isotropic), ratio

G (dB) = $10 \log G$

T_{op} = receiving system operating noise temperature, K

T_{op} (dB) = $10 \log T_{op}$ (relative to $T_{op} = 1$ K)

In 1960, DSN 26-m antennas with 30-percent efficiency at 960 MHz and 1500-K receivers had a figure of merit, G/T_{op} , of 11 dB [2, pp. 2-15]. This was increased to 36 dB at 2.3 GHz in 1965. For the 1980 Voyager-Saturn encounter, DSN 64-m antennas had achieved an efficiency of 50 percent at 8.4 GHz and a G/T_{op} of 58 dB. Since antenna gain G is proportional to the product of effective antenna area and frequency squared, much of the improvement in G/T_{op} is obtained by increasing the operating frequency. Further improvements are obtained by increasing the antenna area and efficiency and by lowering the receiving system noise temperature. The benefit of higher frequencies is limited by technical difficulties such as more stringent antenna surface and pointing tolerances, low-noise amplifier complexity, and, in particular, increased atmospheric losses.

Present DSN 70-m antennas with 69-percent efficiency and microwave receivers at 8.4 GHz achieve a G/T_{op} of 60.4 dB. This is expected to be improved to 68.7 dB in 10 years at 32 GHz; this does not include a planet in the field of view (FOV). The 32-GHz-receiving-antenna estimates in this article apply to individual antennas located at Goldstone, at a 30-deg elevation angle with 90-percent weather confidence. Many future missions to the outer planets are planned as orbiters, some with landers. Therefore, the effects of a planet in the DSN antenna's FOV are important. Saturn is used throughout this article as an example of a planet in the FOV.

The noise temperature contribution of Saturn to T_{op} is estimated at 1.7 and 18 K for the 70-m antenna at 8.4 and 32 GHz, respectively. G/T_{op} alone can be used to compare the TSP of systems, assuming the same transmitter antenna effective area, transmitter power, and modulation/coding performance. G/T_{op} alone is not sufficient, however, when comparing systems with different spacecraft transmitter parameters. A telecommunication system performance ratio (TSP_r) equation is given in this article to compare future microwave and higher frequency

end-to-end systems with the current DSN microwave systems. This article describes present and future deep space TSP estimates for X-band (8.4 GHz) and Ka-band (32 GHz) in the DSN. This follows earlier DSN articles [3] and studies [4].

II. Microwave Receiving System Figure of Merit

In a microwave deep space receiving system, thermal noise is usually the dominant source of noise that limits the system performance. In terms of the receiving system output noise power, the receiving system operating noise temperature is given by [5, p. 7-1]

$$T_{op} = N_o/kBG \quad (2)$$

where

N_o = output thermal noise power of receiving system, W

k = Boltzmann's constant, 1.381×10^{-23} J/K

B = bandwidth, Hz

G = receiver gain, ratio

The receiving system operating noise temperature² is defined at the receiving system input [7]

$$T_{op} = T_a + T_e \quad (3)$$

where

T_a = effective noise temperature of the antenna, K

T_e = effective noise temperature of the receiver, K

The primary noise temperature contributions for the DSN microwave receiving systems are given in Tables 1

² A high-frequency correction [6] $T_c = -0.024f$ (GHz) + $0.000192f^2$ (GHz)²/T + ... is required for the system thermal noise temperature. This is applied in the noise temperature tables of this article as a reduction of the 2.76-K cosmic background noise. For the microwave parameters in this article, the second term is less than 0.02 K. Therefore, only the first term (0.2 and 0.8 K for 8.4 and 32 GHz, respectively), which is independent of T , is used in this article. Further reduction in the cosmic background temperature defined at the receiver input occurs due to atmospheric attenuation. Quantum noise, $T_q = hf/k = 0.0480f$ (GHz) = 0.40 and 1.54 K at 8.4 and 32 GHz, respectively, is included as part of the system noise temperature in these tables.

and 2 for 8.4 and 32 GHz, respectively. The 8.4-GHz systems are operational in the DSN. The 32-GHz low-noise receiver and antenna components have been tested in R&D laboratory and field configurations to provide achievable performance estimates.

The noise temperature contribution of the troposphere is defined for an antenna elevation angle of 30 deg and 90-percent weather confidence for Goldstone.³ At X-band and above, water vapor, precipitation, and clouds cause increased atmospheric attenuation and noise temperature.

Weather confidence is the percentage of time that the predicted TSP is equaled or exceeded due to weather effects only. Equivalently, weather confidence is equal to 100 percent minus the percentage of time the link performance is degraded below the predicted value.

Present DSN antennas use room temperature feed components and 4.5-K physical cryogenically cooled maser amplifiers on the tipping structure of the antenna, where space and accessibility are limited. In the future, primary improvements will occur with use of BWG antennas that facilitate the use of cryogenic cooling for waveguide and feed components below 20 K and maser amplifiers to 1.7 K physical, with resultant lower noise-temperature receiving systems. Combined with antenna gain values, these results are shown in Tables 3 and 4 for G/T_{op} estimates of DSN microwave 34-m and 70-m antennas operating at 8.4 and 32 GHz, now and in the future. Microwave components for the 32-GHz systems have been developed, and preliminary measurements provide data for the estimates shown in Table 4.

As shown in Table 3, the present G/T_{op} performance of 60.1 dB and 60.4 dB (with and without Saturn in the FOV) for the Goldstone 70-m antenna at 8.4 GHz, 30-deg elevation angle, and 90-percent weather confidence is used as a baseline for this study.

Antenna efficiency improvements are planned to maximize the 32-GHz antenna gain, in addition to 8.4- and 32-GHz noise temperature reductions. Table 4 shows 68.7 dB is predicted as the best microwave G/T_{op} value. This is achieved at 32 GHz on a DSN 70-m BWG antenna with the feed and waveguide components cooled to 20 K and the maser to 1.7 K physical. This represents an 8.3-dB

improvement relative to the present baseline performance of 60.4 dB at 8.4 GHz.

The first deep space test application of 32 GHz will use DSS 13 and the Mars Observer (MO) spacecraft with the Ka-band link experiment (KABLE).⁴ MO will have a simultaneous 8.4- and 33.7-GHz transmit capability. This provides an opportunity to verify the end-to-end Ka-band telemetry system performance by direct comparison with X-band.

III. Telecommunication System Performance

The Friis free-space received-power transmission formula [8] is manipulated and expanded to a form useful for computing comparison ratios of one telecommunication system relative to another

$$TSP_r = Pr \times ATr \times Mr \quad (4)$$

where

TSP_r = telecommunication system performance comparison ratio

$$TSP_r \text{ (dB)} = Pr \text{ (dB)} + ATr \text{ (dB)} + Mr \text{ (dB)}$$

Pr = transmitter power comparison ratio

$$Pr \text{ (dB)} = 10 \log Pr$$

ATr = transmitter antenna effective area comparison ratio

$$ATr \text{ (dB)} = 10 \log ATr$$

Mr = receiving system figure of merit G/T_{op} comparison ratio

$$Mr \text{ (dB)} = 10 \log Mr$$

The form of the power transmission formula used in Eq. (4) is useful for a transmitter with a size-limited antenna [9]. For most of this article, microwave spacecraft transmitter power and antenna area (and efficiency) are assumed unchanged with frequency, that is, $Pr \times ATr = 1$. For this assumption, Eq. (4) shows that the system figure of merit ratio, $Mr = (G/T_{op})_r$, is a useful and convenient measure for comparing relative performances.

³ S. Slobin, "DSN Telecommunications Interfaces, Atmospheric and Environmental Effects," TCI 40, Rev. C (internal document), to be published in *DSN/Flight Project Interface Design Document*, JPL 810-5, Rev. D, Jet Propulsion Laboratory, Pasadena, California, 1992.

⁴ S. Butman and J. Meeker, *DSN Advanced Systems—Mars Observer Ka Band Link Experiment Plan*, JPL D-8799 (internal document), Jet Propulsion Laboratory, Pasadena, California, August 1991.

The reference telecommunications system for this article is the Voyager spacecraft's 8.4-GHz transmitting microwave system and the DSN Goldstone 70-m receiving station. The Voyager spacecraft has a 3.7-m-diameter antenna and approximately 20-W X-band transmitter output power. The antenna has 62-percent area efficiency, about 0.6-deg half-power beamwidth, and a pointing accuracy of about 0.1 deg.

An exception to the use of the reference system is included as an example where $Pr \times ATr \neq 1$, emphasizing the need to compare the end-to-end TSP. Future use of higher frequencies with millimeter and optical wavelengths for deep space communications is expected to allow the use of smaller spacecraft transmitting antennas with less transmitter power. A future study summarizing and comparing the performance of higher frequency systems with the current microwave systems is planned.

For this article, the example where $Pr \times ATr \neq 1$ assumes a 32-GHz spacecraft transmitting system with a shaped, clear aperture, BWG antenna, and a high-efficiency 30-W transmitter (see Table 5). The antenna diameter is 3.7 m, but the effective antenna area is 1.4 times greater (+1.4 dB) than the reference system because the area efficiency is 86 percent. The 30-W transmitter provides 50 percent more power (+1.8 dB) than the reference system, and $Pr \times ATr = 2.1$ (+3.2 dB).

Table 5 shows an estimated 8.3-dB-improved TSP in 10 years for ground-based, 32-GHz microwave systems compared with the DSN Goldstone 8.4-GHz 70-m antenna's present 60.4-dB G/T_{op} performance. A 1988 study gives similar results.⁵ For the portion of the study where $Pr \times ATr = 1$ (no transmitter improvements), TSP_r is determined by the relative figure of merit, Mr .

IV. Tolerances

The values in the text and tables do not indicate tolerances. The DSN receiving stations' X-band G/T_{op} tolerance is ± 0.4 dB. The TSP at 30-deg antenna elevation angle and 90-percent weather confidence is estimated as ± 0.6 dB. The 32-GHz G/T_{op} tolerance uncertainties are much worse than the 8.4-GHz tolerances because weather effects dominate the ground station performance, and there is no deep space experience to verify existing models. A ± 2 -dB tolerance is estimated for 32-GHz end-to-end performance. A precision of 0.1 dB was included in the text

and tables to avoid accumulation of round-off errors; this does not imply accuracy to 0.1 dB.

V. Conclusion and Future Directions

The data presented for the DSN ground system point to significantly improved figure of merit and telecommunications system performance in the coming years. Increasing the operating frequency from the present 8.4 GHz to 32 GHz is expected to result in a net gain of 8.3 dB (without significant planet noise) and 7.0 dB (with Saturn in the FOV) in the next 10 years, relative to the current DSN Goldstone 70-m antenna.

Use of a high-efficiency 32-GHz spacecraft transmit system can increase telecommunications system performance another 3.2 dB, increasing the net performance gain from 8.3 to 11.5 dB without a planet in the FOV.

The potential improvements for the 32-GHz systems are important for future deep space communications. These improvements could provide telecommunication systems with higher data rates for the spacecraft-to-Earth link and/or some combination of lower transmitter power, smaller antenna size, and lower weight. The choice of which improvements to use, and when, will depend upon practical considerations that include technical readiness and cost.

Future options for performance improvements in DSN telecommunications systems during the time beyond 2001 include:

- (1) Higher receiving-station antenna gain resulting from increased antenna-aperture areas, probably using arrays. Lower system-noise temperature using antenna BWG technology with super-cooled, low-noise maser amplifiers and cooled waveguide components.
- (2) Use of wavelengths/frequencies that better optimize the combination of receiving station G/T_{op} , spacecraft transmitter power, spacecraft antenna size, and receiver performance according to constraints affecting deep space missions.
- (3) Improved spacecraft transmitter components, resulting in higher antenna efficiencies and power.

Future use of Earth-orbiting deep space relay satellites will provide a major change for the space-to-ground telecommunications system. This architecture eliminates tropospheric loss and noise and potentially allows greatly increased data rates for telemetry.

⁵ J. W. Layland, R. C. Clauss, R. L. Horttor, D. J. Mudgway, R. J. Wallace, and J. H. Wilcher, *Ka-band Study—1988, Final Report*, JPL D-6015 (internal document), Jet Propulsion Laboratory, Pasadena, California, February 15, 1989.

Acknowledgments

A. Cha and D. Bathker supplied the microwave antenna efficiency estimates. S. Slobin supplied the troposphere microwave parameters, information about Saturn's contribution to the system noise temperature, and accuracy estimates.

References

- [1] D. Schuster, C. T. Stelzried, and G. S. Levy, *The Determination of Noise Temperatures of Large Antennas*, JPL Technical Report No. 32-97, Jet Propulsion Laboratory, Pasadena, California, May 1, 1961.
- [2] N. A. Renzetti, C. T. Stelzried, G. K. Noreen, S. D. Slobin, S. M. Petty, D. L. Trowbridge, H. Donnelly, P. W. Kinman, J. W. Armstrong, N. A. Burow, M. K. Tam, J. W. Layland, and A. L. Berman, *The Deep Space Network—A Radio Communications Instrument for Deep Space Exploration*, JPL Publication 82-104, Jet Propulsion Laboratory, Pasadena, California, July 15, 1983.
- [3] M. S. Reid, R. C. Clauss, D. A. Bathker, and C. T. Stelzried, "Low-Noise Microwave Receiving Systems in a Worldwide Network of Large Antennas," *Proceedings of the IEEE*, pp. 1330-1335, September 1973.
- [4] P. D. Potter, M. S. Shumate, C. T. Stelzried, and W. H. Wells, *A Study of Weather-Dependent Data Links for Deep Space Applications*, JPL Technical Report No. 32-1392, Jet Propulsion Laboratory, Pasadena, California, October 15, 1969.
- [5] C. Stelzried, *The Deep Space Network—Noise Temperature Concepts, Measurements, and Performance*, JPL Publication 82-33, Jet Propulsion Laboratory, Pasadena, California, September 15, 1982.
- [6] M. Viggh, "How Noisy Is That Load," *Microwaves*, vol. 15, no. 1, pp. 54, 56, January 1976.
- [7] "IRE Standards on Electron Tubes: Definitions of Terms, 1962 (62 IRE 7.S2)," *Proceedings of the IEEE*, pp. 434-442, March 1963.
- [8] H. T. Friis, "A Note on a Simple Transmission Formula," *Proceedings of the IRE*, pp. 254-256, 1946.
- [9] K. W. Linnes, W. D. Merrick, and R. Stevens, "Ground Antenna for Space Communication System," *IRE Transactions on Space Electronics and Telemetry*, p. 45, March 1960.

Table 1. DSN receiving systems' noise temperature performance at 8.4 GHz, K.

Configuration	Antenna noise temperature contributions, K							Maser ^b T_e , K	Total T_{op} , K	
	Cosmic background ^a	Atmosphere	Antenna	BWG	Dichroic plate	Feed and WG	T_a , K		SFOV ^c	
Now ^{c,d}	2.6	5.2	4.3	0	1.2	6	19.3	4	23.3	25.0
10 years ^{c,e}	2.6	5.2	4.3	1	1.2	0.3	14.6	1	15.6	17.3

Conditions: 30-deg antenna elevation angle and Goldstone, California, 90-percent weather confidence.

^a Includes high-frequency thermal noise temperature correction, T_c .

^b Includes quantum noise temperature, T_q , and noise contribution from the follow-on amplifiers.

^c Saturn in the field of view (SFOV) adds about 1.7 K to T_{op} at 8.4 GHz.

^d Non-BWG, uncooled feed, maser at 4.5 K, physical temperature.

^e BWG, feed cooled, maser at 1.7 K, physical temperature.

Table 2. DSN receiving systems' estimated noise temperature performance at 32 GHz, K.

Configuration	Antenna noise temperature contributions, K							Maser ^b T_e , K	Total T_{op} , K	
	Cosmic background ^a	Atmosphere	Antenna	BWG	Dichroic plate	Feed and WG	T_a , K		SFOV ^c	
Now ^{c,d}	1.8	25.4	5	0	3	7	42.2	11	53.2	71.2
10 years ^{c,e}	1.8	25.4	5	1.7	2	0.5	36.4	4	40.4	58.4

Conditions: 30-deg antenna elevation angle and Goldstone, California, 90-percent weather confidence.

^a Includes high-frequency thermal noise temperature correction, T_c .

^b Includes quantum noise temperature, T_q , and noise condition from the follow-on amplifiers.

^c SFOV adds 18 K to T_{op} at 32 GHz.

^d Achievable for R&D configuration non-BWG, uncooled feed, maser at 4.5 K, physical temperature.

^e BWG, feed cooled, maser at 1.7 K, physical temperature.

Table 3. DSN 8.4-GHz receiving antennas' G/T_{op} performance.

Configuration	Diam., m	Antenna gain, dB	T_{op} , K	G/T_{op} , dB	
				SFOV ^a	
Now	34	68.2	23.3	54.5	
10 years	34	68.5	15.6	56.5	
Now	70	74.1	23.3	60.4 ^b	60.1 ^b
10 years	70	74.7	15.6	62.8	62.4

Conditions: 30-deg antenna elevation angle and Goldstone, California, 90-percent weather confidence, resulting in 0.08-dB atmospheric loss.

Ten-year improvements: BWG with cooled feeds and masers cooled to 1.7 K, physical temperature.

^a Saturn in the field of view (SFOV).

^b Baseline performances.

Table 4. DSN 32-GHz receiving antennas' G/T_{op} performance.

Configuration	Diam., m	Antenna gain, dB	T_{op} , K	G/T_{op} , dB	
				SFOV ^a	
Now	34	77.2	53.2	60.0	
10 years	34	79.2	40.4	63.1	
Now	70	82.4	53.2	65.2	63.9
10 years	70	84.8	40.4	68.7	67.1

Conditions: 30-deg antenna elevation angle and Goldstone, California, 90-percent weather confidence, resulting in 0.42-dB atmospheric loss.

Now: Achievable based on DSS-13 R&D 34-m antenna gain measurements and Table 2 noise temperature estimates.

Ten-year improvements: Antenna gains, BWG with cooled feeds, and masers at 1.7 K, physical temperature.

^a Saturn in the field of view (SFOV).

Table 5. DSN 10-year future microwave TSP comparison, decibels relative to present 70-m antenna at 8.4 GHz with 60.4-dB figure of merit, G/T_{op} without SFOV and 60.1 dB with SFOV.

Parameter ratio	8.4 GHz			32 GHz			32-GHz HE spacecraft TS		
	34 m	70 m	70-m, SFOV ^a	34 m	70 m	70-m, SFOV ^a	34 m	70 m	70-m, SFOV ^a
$PrAr$ (transmitter power \times antenna area)	0	0	0	0	0	0	3.2	3.2	3.2
Mr (receiving system figure of merit)	-3.9	2.4	2.3	2.7	8.3	7.0	2.7	8.3	7.0
TSP_r (telecommunications system performance)	-3.9	2.4	2.3	2.7	8.3	7.0	5.9	11.5	10.2

Conditions:

Transmitter: Same power and antenna efficiencies except for 32-GHz high-efficiency spacecraft transmit system (HE spacecraft TS) with 3.7-m antenna diameter.

Receiving: 30-deg antenna elevation angle and Goldstone, California, 90-percent weather confidence.

^a Saturn in the field of view (SFOV).

P-15
N92-24322

JJ 57445

Position Determination of a Lander and Rover at Mars With Earth-Based Differential Tracking

R. D. Kahn, W. M. Folkner, and C. D. Edwards
Tracking Systems and Applications Section

A. Vijayaraghavan
Navigation Systems Section

The presence of two or more landed or orbiting spacecraft at a planet provides the opportunity to perform extremely accurate Earth-based navigation by simultaneously acquiring Doppler data and either Same-Beam Interferometry (SBI) or ranging data. Covariance analyses were performed to investigate the accuracy with which lander and rover positions on the surface of Mars can be determined. Simultaneous acquisition of Doppler and ranging data from a lander and rover over two or more days enables determination of all components of their relative position to under 20 m. Acquiring one hour of Doppler and SBI enables three-dimensional lander-rover relative position determination to better than 5 m. Twelve hours of Doppler and either SBI or ranging from a lander and a low-circular or half-synchronous circular Mars orbiter makes possible lander absolute-position determination to tens of meters.

I. Introduction

The planet Mars will undergo an intensive program of unmanned exploration in coming years. In 1993, the Mars Observer spacecraft will begin to map the Martian surface. The Commonwealth of Independent States (former Soviet Union) is scheduled to launch missions to Mars in 1994 and 1996, in which a number of landers and balloons will be deployed. The U.S. Space Exploration Initiative encompasses a broad range of unmanned and manned trips to the Moon and Mars, which will involve a variety of orbiters, landers, and rovers. The presence of multiple spacecraft at Mars could provide the opportunity to perform ex-

tremely accurate Earth-based navigation by using a radio metric technique called Same-Beam Interferometry (SBI) [1,2,3,4]. SBI measurements provide information about spacecraft-spacecraft separation in the plane perpendicular to the Earth-spacecraft line of sight, naturally complementing the position and velocity information obtained from Doppler and ranging measurements. SBI could prove to be a valuable navigational tool supporting future Mars missions, potentially aiding in the final hours of Mars approach navigation, enabling improvement of Mars orbiter ephemerides and providing measurements of the relative positions of landers and rovers with an accuracy of several meters.

Radio metric data types currently used by the DSN for performing navigation of interplanetary spacecraft include ranging, Doppler, and delta very long baseline interferometry (Δ VLBI). While the ranging and Doppler systems provide a direct measurement of line-of sight spacecraft range and range-rate, Δ VLBI provides a measure of spacecraft angular position. The current operational DSN Δ VLBI navigation system can provide a measurement of spacecraft angular position, relative to a fixed extragalactic radio source, with an accuracy of about 30 nrad (equivalent to 4 km at a distance of 1 AU). SBI is a variant of Δ VLBI, with a measurement precision of tens of picoradians (equivalent to several meters at a distance of 1 AU). SBI involves observation of two (or more) spacecraft that are so angularly close that they simultaneously lie within the beamwidth of Earth-based antennas. Simultaneous observation of multiple spacecraft enables orders of magnitude of improvement in accuracy because of tremendous error cancellation resulting when spacecraft observables are differenced.

When SBI and Doppler are simultaneously acquired from two landed spacecraft (e.g., a lander and a rover), it is possible to obtain a several-meter-level determination of the relative position of the two spacecraft on the Martian surface. While traversing the Martian landscape, a rover will need to rely heavily on in-situ navigation techniques, possibly including the use of radio links to Mars-orbiting spacecraft. Earth-based Doppler and SBI provide an independent means of establishing the lander-rover relative position at the several-meter level in three dimensions. SBI could serve as a complement to in-situ navigation, or be used to periodically calibrate lander-rover position measurements obtained from local navigation techniques.

Combining SBI and Doppler is not the only means of determining the position of a rover relative to a lander with Earth-based measurements. Simultaneous Doppler and ranging acquired from both spacecraft at one or more DSN antennas also enables accurate determination of spacecraft-spacecraft relative position. A number of Doppler-ranging examples are examined in this study.

Absolute lander position in the Mars reference frame can be established by simultaneously acquiring Doppler and SBI data or Doppler and range data from the lander and a Mars orbiter. The Doppler data are able to determine the lander components in the plane perpendicular to the Mars spin axis; lander-orbiter SBI, or simultaneous Earth-based ranging to the lander and orbiter, constrains the third component of lander position since the orbiter trajectory is tied to the Mars center of mass.

II. SBI Description

SBI provides a measure of spacecraft-spacecraft angular separation as viewed from Earth. Consider two spacecraft at Mars simultaneously transmitting radio signals to two Earth antennas. Because the two spacecraft are very close angularly (the diameter of Mars as viewed from Earth is at most 124 μ rad), each Earth antenna can simultaneously acquire the two spacecraft signals (Fig. 1). The signal from a given spacecraft does not arrive at each of the two Earth antennas simultaneously; the delay in signal reception depends on the angle between the vector connecting the two Earth antennas and the vector from the Earth to the spacecraft. However, the measured delay is corrupted by a variety of error sources, including delays caused by system instrumentation and by signal propagation through neutral and charged media. These errors largely cancel when observables obtained simultaneously from two spacecraft are differenced. The resulting double-differenced delay provides an extremely precise measurement of spacecraft-spacecraft angular separation in the direction parallel to the plane-of-sky projection of the Earth baseline.

SBI observables involving two spacecraft and two Earth antennas are obtained by accumulating spacecraft signal phase at each of the two Earth antennas. If the signal transmitted by each spacecraft has frequency f , then the double-differenced delay τ may be expressed as

$$\tau(t) = \frac{(\phi_{12}(t) - \phi_{11}(t)) - (\phi_{22}(t) - \phi_{21}(t)) + b}{f}$$

where $\phi_{ij}(t)$ represents the phase of the signal transmitted from spacecraft i and received at station j , and b is an unknown integer. Because of the unknown integer bias, an SBI measurement does not directly provide the double-differenced delay. However, SBI measurements obtained continuously over a time interval share the same bias and thus provide a precise measure of temporal changes in the double-differenced delay. Given sufficient a priori information about the spacecraft states, the integer cycle ambiguity can be resolved and the absolute double-differenced delay determined. Typically, a priori information about the spacecraft states is not sufficient to determine the integer-cycle ambiguity, and the SBI phase bias must therefore be estimated. If the sigma on the SBI phase-bias estimate is smaller than 1/6 of a cycle of carrier phase, the integer-cycle ambiguity can be resolved with 99-percent confidence.

It is worth observing that there is an alternative scheme for resolving the SBI phase bias in a relatively short track-

ing arc. This strategy requires that the spacecraft transmit a spectrum of several tones, rather than a single carrier. Preliminary SBI group-delay observables are formed from pairs of tones that are spaced by successively wider intervals. For the narrowest tone spacing, the integer phase ambiguity can be resolved from a priori knowledge. If the frequencies of the tone pairs are carefully selected, then once the delay is resolved for a given pair, the accuracy of the resulting delay measurement is sufficient to enable resolution of the integer-cycle ambiguity for the next (more widely spaced) tone pair. One can bootstrap up through a series of increasingly wider spaced tone pairs, until the SBI phase bias for the carrier is resolvable. This technique could enable SBI carrier-phase bias resolution over time scales on the order of minutes.

III. Lander–Rover SBI Error Budget

SBI measurements are corrupted by a variety of sources, including thermal noise in the ground receiver, signal propagation through neutral and charged media, and nonlinearities in the phase response of station instrumentation. The effect of these errors on SBI measurement accuracy is described in the following error budget, which is similar to several previously developed SBI error budgets [1,2,3,4]. Here the SBI error is given in millimeters. The equivalent angular error in spacecraft–spacecraft relative position is obtained by dividing the delay error by the length (in millimeters) of the plane-of-sky projection of the vector connecting the two Earth antennas. If the two Earth antennas are located at the DSN complexes in Goldstone and Canberra, the length of this projection is typically 8000–10,000 km.

The SBI measurement error is computed for three different cases, which are summarized in Table 1. In Case 1, the lander and rover transmit a 2.3-GHz (S-band) signal with omnidirectional antennas. Signal power is 5 W. In Case 2, the spacecraft have 0.5-m diameter directional antennas, and the transmitted signal is 8.4 GHz (X-band), with 10 W of power. Case 3 is identical to Case 2, except that the spacecraft transmit at 32.5 GHz (Ka-band). In all three cases, it is assumed that 34-m DSN antennas are used for reception.

A. System Noise

The system noise error depends on the ratio of received signal power to the noise power generated in the ground receiver. The 1-sec voltage signal-to-noise ratio (SNR_v) has been calculated for each of the three cases listed in Table 1, assuming an Earth–Mars distance of 2.5 AU (the maximum possible), and system noise temperature of 20 K

at 2.3 GHz, 25 K at 8.4 GHz, and 80 K at 32.5 GHz. Antenna efficiency is 0.7. Given these assumptions, $\text{SNR}_v = 3.6$ for Case 1, $\text{SNR}_v = 152$ for Case 2, and $\text{SNR}_v = 361$ for Case 3. The SBI system noise error is given by

$$\epsilon = \frac{2\lambda}{2\pi\text{SNR}_v\sqrt{T}} \text{ mm}$$

where λ is the wavelength of the signal in millimeters, and T is the integration time in seconds. The 2 in the numerator accounts for the fact that a separate error occurs for each of the four received signals. For an integration time of 5 minutes,

$$\epsilon = 0.664 \text{ mm} \quad (\text{Case 1})$$

$$\epsilon = 0.004 \text{ mm} \quad (\text{Case 2})$$

$$\epsilon = 0.0005 \text{ mm} \quad (\text{Case 3})$$

B. Instrumental Phase Dispersion

Uncalibrated phase shifts occurring in station instrumentation produce delays that vary with signal frequency. It should be possible to calibrate these delays to the level of 0.5 deg by using digital circuitry, such as that used in the DSN narrow-channel-bandwidth VLBI System [5]. The SBI error due to instrumental phase dispersion can be computed as

$$\epsilon = 2 \times \left(\frac{0.5}{360} \right) \times \lambda \text{ mm}$$

where λ is the wavelength of the signal in millimeters, and the factor of 2 accounts for the fact that a separate error occurs for each of the four signal paths. This expression yields

$$\epsilon = 0.361 \text{ mm} \quad (\text{Case 1})$$

$$\epsilon = 0.100 \text{ mm} \quad (\text{Case 2})$$

$$\epsilon = 0.026 \text{ mm} \quad (\text{Case 3})$$

C. Troposphere

The time delay due to signal transit through the tropospheres of Earth and Mars is independent of signal frequency. The SBI error resulting from uncertainty in the zenith delay at the Earth stations is given by

$$\epsilon = \sqrt{2}\Delta\theta \times 40 \text{ mm} \times \cos(E)/\sin^2(E)$$

where 40 mm is the zenith-troposphere delay uncertainty (after calibration), $\Delta\theta$ (radians) is the difference in lander

and rover elevation as viewed from Earth, and E is the elevation of Mars as viewed from Earth. The factor of $\sqrt{2}$ accounts for an independent error at each Earth station. Assuming an Earth-Mars distance of 2.5 AU, lander-rover separation of 100 km, and elevation of 15 deg, the Earth troposphere delay error is $\epsilon = 0.0002$ mm.

Water vapor fluctuations in the Earth's troposphere contribute an SBI error of the form [4]:

$$\epsilon = \sqrt{2} \times 14.4 \times \Delta\theta^{5/6} \text{ mm}$$

For the lander-rover cases considered here, the contribution of water vapor fluctuations to the SBI error is 7×10^{-5} mm.

Viking lander measurements and Viking orbiter radio-occultation measurements of the Mars atmosphere indicate that atmospheric pressure at the Martian surface is less than 10 mb, and that the atmospheric mean molecular weight is 43 [6,7]. To obtain an estimate of the total (un-calibrated) delay through the Mars troposphere, it is assumed that the forces of pressure and gravity on each slice of atmosphere are exactly balanced ($dp/dh = -\rho g$, where p is pressure, h is altitude, ρ is density, and $g \approx 3.7 \text{ m/sec}^2$ is the acceleration due to gravity at the surface of Mars) and that the atmosphere is an isothermal ideal gas with a constant mixing ratio (which implies that $p/p_0 = \rho/\rho_0$, where p_0 and ρ_0 are the atmospheric surface pressure and density). These assumptions imply a total zenith dry troposphere delay on the order of 40 mm [8]. The Mars troposphere delay error is written as

$$\epsilon = 0.5\Delta\gamma \times 40 \text{ mm} \times \cos(e)/\sin^2(e)$$

where $\Delta\gamma$ is the difference in Earth station elevation (in radians) as viewed from Mars, and e is the elevation of the Earth as viewed from Mars. The factor of 0.5 accounts for the fact that the Mars troposphere delay error is highly correlated between spacecraft (measurements of the Earth's troposphere indicate density correlations of 90 percent over hundreds of kilometers [9]). For an Earth-Mars distance of 2.5 AU, Earth station separation of 10,000 km, and Earth elevation of 15 deg, the Mars troposphere delay error is $\epsilon = 0.008$ mm.

Viking lander measurements indicate that columnar water-vapor content in the Mars atmosphere is three orders of magnitude smaller than on Earth; Mars water-vapor fluctuations should not contribute appreciably to the SBI measurement error.

D. Ionosphere

The ionospheres of Earth and Mars introduce interferometric delay measurement errors that are frequency dependent. Measurement of total electron content (TEC) along the line of sight between Earth antennas and Global Positioning System (GPS) satellites can be mapped to the line of sight from Earth antennas to spacecraft at Mars, which provides TEC calibrations at the level of 5×10^{16} electrons/m². The SBI error after calibration is expressed as

$$\epsilon = \sqrt{2}\Delta\theta \times \frac{2233}{\nu^2} \times 5 \text{ mm}$$

Here, the factor of 5 is an upper bound on the derivative of the mapping function, $\Delta\theta$ is the angular separation of the spacecraft as viewed from Earth, and ν is the signal frequency in gigahertz.

Temporal fluctuations in the electron content along the signal through the Earth's ionosphere result in an SBI error of the form [4]:

$$\epsilon = \sqrt{2}\Delta\theta^{5/6} \times \frac{3510}{\nu^2} \text{ mm}$$

For the three cases under consideration here, the total SBI error due to the Earth ionosphere is:

$$\begin{aligned} \epsilon &= 0.003 \text{ mm} && \text{(Case 1)} \\ \epsilon &= 0.0002 \text{ mm} && \text{(Case 2)} \\ \epsilon &= 0.00001 \text{ mm} && \text{(Case 3)} \end{aligned}$$

A reasonable estimate of the total zenith Mars ionospheric delay for an S-band signal is 250 mm [6]. The SBI error due to the Mars ionosphere can be written as:

$$\epsilon = 0.5\Delta\gamma \times \frac{250 \text{ mm}}{(\nu/2.3)^2} \times \cos(e)/\sin^2(e)$$

with e and $\Delta\gamma$ as defined above. Here, a $1/\sin(e)$ ionospheric mapping function is assumed. The factor of 0.5 accounts for correlations between the ionospheric delay on the measurements from each of the spacecraft. The resulting SBI error is:

$$\begin{aligned} \epsilon &= 0.05 \text{ mm} && \text{(Case 1)} \\ \epsilon &= 0.004 \text{ mm} && \text{(Case 2)} \\ \epsilon &= 0.0002 \text{ mm} && \text{(Case 3)} \end{aligned}$$

SBI errors due to temporal fluctuations in the Mars ionosphere are small, and not considered here.

E. Station Locations and Universal Time/Polar Motion

Uncertainties in Earth station locations and Earth orientation lead to an SBI error of the form, $\varepsilon = \Delta\theta\varepsilon_{SL}$, where ε_{SL} represents the error in relative positions of the Earth stations, and $\Delta\theta$ is the angular separation (in radians) of the two spacecraft as viewed from Earth. The combined station location and Universal Time/Polar Motion (UTPM) uncertainty is taken to be 70 mm, assuming real-time UTPM estimates based on daily GPS measurements combined with VLBI observations [10]. The resulting SBI error is $\varepsilon = 0.00002$ mm.

F. Solar Plasma

An SBI observable is formed by doubly differencing phase measurements made along the four lines of sight connecting two spacecraft and two Earth antennas (Fig. 1). The four signal paths are separated by hundreds or thousands of kilometers while traversing interplanetary space, which results in imperfect cancellation of the delay error induced by charged particles in the solar wind. The solar plasma delay error is inversely proportional to the square of signal frequency and increases as the Sun–Earth–spacecraft angle decreases. A thin-screen frozen turbulence model is used to model the plasma-induced error [11]. For a lander and rover separated by 100 km, and a Sun–Earth–Mars angle of 20 deg, the plasma-induced SBI error is:

$$\varepsilon = 0.059 \text{ mm} \quad (\text{Case 1})$$

$$\varepsilon = 0.004 \text{ mm} \quad (\text{Case 2})$$

$$\varepsilon = 0.0003 \text{ mm} \quad (\text{Case 3})$$

The SBI measurement errors are summarized in Table 2. For Case 1, the measurement error is dominated by system noise, while instrumental phase dispersion is the principal error for Cases 2 and 3. The 0.76-mm measurement error in Case 1 is equivalent to an angular position error of 95 prad, or 14 m in spacecraft–spacecraft separation for spacecraft at a distance of 1 AU (the Earth–Mars distance varies from 0.5–2.5 AU). For Cases 2 and 3, the SBI measurement errors are equivalent to spacecraft–spacecraft separation errors of 2 m and 0.6 m at 1 AU.

IV. Lander–Rover Relative Position Determination

A. Lander–Rover Positioning With SBI and Doppler Data

Covariance analyses were performed to determine the accuracy with which the relative positions of a Mars lander and rover can be determined by using Doppler and

SBI data. The lander is located at 0 deg longitude and +30 deg latitude, and the rover is 100 km north of the lander. The Doppler–SBI data-acquisition period ranges from 1–3.75 hr. Analyses were conducted for each of the three radio frequency configurations listed in Table 1.

The simulated data are acquired for August 15, 1994. Lander and rover Doppler data are acquired at Goldstone; SBI data are acquired along the Goldstone–Canberra baseline. All radio links are assumed to be two-way. For Case 1 (2.3-GHz carrier frequency), the 5-min Doppler are weighted at 0.45 mm/sec; for Cases 2 and 3 (8.4-GHz carrier or 32.5-GHz carrier), the 5-min data are weighted at 0.045 mm/sec. The SBI data are weighted in accordance with the error budget presented above.

A summary of the spacecraft estimated and considered (unadjusted) parameters appears in Table 3. Mars UT1–UTC, the orientation of the Mars pole in inertial space, and the ephemerides of Earth and Mars are estimated together with the lander and rover positions. The a priori sigma for Mars pole orientation is 50 μ rad in right ascension and declination [12], and the a priori uncertainty of Mars UT1–UTC is a conservative 30 μ rad [14]. Considered parameters include Earth and Mars zenith troposphere delays, Earth and Mars zenith TEC, and Earth-station locations. The media sigmas for Earth represent post-calibration uncertainties; those for Mars represent total uncertainties.

As noted above, the SBI-phase observable contains an integer bias that must (initially) be estimated together with the lander and rover positions. If the integer-cycle ambiguity can be resolved, the full strength of the data can be used to estimate the spacecraft positions. Preliminary covariance runs indicated that the integer bias can be unambiguously determined, even for data arcs as brief as 1 hr. For Cases 2 and 3, it was found that fixing the bias in a 3.75-hr data arc marginally improves the lander–rover relative-position uncertainty; for Case 1, results are improved by a factor of 3. For all three cases, fixing the SBI bias in a 1- or 2-hr data arc improves the lander–rover relative-position uncertainty by up to an order of magnitude. All the results appearing in this article were obtained with the SBI bias fixed.

Figures 2–4 present covariance analysis results for a variety of representative cases. In each of these figures, the lander–rover relative-position error is plotted in Mars-centered, Mars-fixed coordinates. The z-direction is parallel to the Mars axis of rotation, the x-direction (“spin radius” direction) points perpendicularly outward from the Mars spin axis through the lander’s position, and the

the y-direction (longitude) is perpendicular to the x- and z-directions.

Figure 2 illustrates the relative position accuracy achievable when Doppler data alone are acquired at Goldstone from both the lander and rover. The Doppler signature induced by Mars rotation enables accurate determination of the components of spacecraft separation that lie in the plane perpendicular to the planet's spin axis. However, Doppler data are insensitive to the third component of lander-rover separation; the z-component error is several kilometers.

When SBI and Doppler data are concurrently acquired, all three components of lander-rover separation can be accurately determined; SBI data complement the Doppler data by providing a precise measurement of spacecraft separation in the plane of the sky. Figure 3 shows the relative position accuracy for the lander and rover when 3.75 hr of SBI and Doppler data are acquired. In Cases 2 and 3 (8.4- or 32.5-GHz carrier frequency), the z-component of lander-rover relative separation can be determined to the several-meter level. Note that the SBI data also improve the determination of the x- and y-components. In Case 1 (2.3-GHz carrier), the SBI and Doppler data are somewhat weaker, enabling x- and y-determination at the level of several meters, and z to about 13 m. The lander-rover relative-position errors depicted in Fig. 3 are the root-sum-square (RSS) of the considered and computed errors; in all cases, the computed error is dominant.

If the transmitting and receiving antennas have 10 deg elevation-visibility cutoffs, the maximum length of time that a Mars lander can be observed along the Goldstone-Canberra baseline is about 3.75 hr. However, spacecraft power constraints or antenna-pointing limitations could further restrict the length of time that Mars landers and rovers are able to transmit to the Earth. Figure 4 illustrates relative position accuracy as a function of observation time. In Cases 2 and 3 (8.4- or 32.5-GHz carrier frequency), acquisition of as little as one hour of data enables determination of all three components of lander-rover separation to the few-meter level. In Case 1 (2.3-GHz carrier frequency), a 1-hr data arc does not provide quite the same level of position accuracy, though it does enable RSS lander-rover relative position determination of under 30 m—more than two orders of magnitude improvement over the position accuracy attainable with a 3.75-hr arc of 8.4-GHz (X-band) Doppler alone.

Additional Doppler-SBI cases were investigated with lander-rover separation oriented East-West, and Northeast-Southwest, and with the lander-rover pair located at 0 deg or -30 deg latitude. It was found that

the ability to determine the relative separation of the two spacecraft does not vary significantly with latitude or rover-lander orientation. The uncertainty in the z-component of the rover-lander separation vector changes by between 1 and 3 m as spacecraft latitude and orientation are varied. This variation in z-uncertainty results from slight differences in sensitivity to Mars pole orientation. Because the separation of the lander and rover was relatively small in the cases studied here, sensitivity of the SBI data to Mars pole orientation was not great enough to permit significant improvement upon the a priori uncertainty. However, measurements involving a network of three or more widely spaced Mars landers could enable determination of Mars pole orientation to the meter level.

B. Lander-Rover Positioning With Ranging and Doppler Data

Another method of determining the relative position of two spacecraft on the Martian surface is to simultaneously acquire Doppler and range data from both spacecraft. To determine the relative position accuracy attainable by using this strategy, lander and rover Doppler and ranging passes were scheduled, with the Doppler weighted at 0.045 mm/sec over 5 min, and the ranging data weighted at 1 m. Range measurements were also assigned a 5-m bias to account for uncalibrated delays in the system hardware.

The ability of ranging to resolve the z-component of rover-lander separation depends to some extent on Earth-Mars geometry, which changes significantly over a period of months. Two cases are considered here: (1) data are acquired for August 15, 1994, from the same lander-rover pair considered in the SBI-Doppler example above; and (2) data are acquired for June 2, 1994, from a lander-rover pair with the lander located at 90 deg East longitude, -30 deg South longitude, and the rover 100 km north of the lander. In the first case, Earth declination as viewed from Mars is 3 deg, so Earth-Mars ranging measurements have little projection in the z-direction. In the second case, the declination of Earth is close to 24 deg. In each case, ranging and Doppler data are acquired at the three DSN antennas in Goldstone, Canberra, and Madrid whenever the spacecraft are visible. The estimation strategy is the same as above (Table 3).

Figure 5 illustrates the z-component uncertainty as a function of scan length for the two cases. In the first case (data acquired beginning August 15, 1994), a single day of tracking (the spacecraft are visible from Earth for approximately 11 hr) enables determination of the z-component of lander-rover separation to about 160 m. Dramatic reduction in the z-component error occurs during the second day of tracking; the geometry has evolved

sufficiently so that uncertainties in Mars pole orientation and planetary ephemerides can be separated from relative z-component uncertainties. After three days of data acquisition, the z-component of spacecraft separation is determined to about 15 m. In the second case (data acquired beginning June 2, 1994), acquisition of a single day of Doppler and ranging determines the z-component of rover-lander separation to the 20-m level; multiple-day observations enable modest improvement in the determination of the z-component.

It should be noted that the ability of Doppler and ranging to perform accurate lander-rover relative position determination may be significantly improved if the Mars pole orientation and Mars UT1 are more tightly constrained. Though these parameters are currently known at the 50- μ rad level, the monitoring of multiple surface beacons at Mars in the future could enable determination at the level of tenths of microradians.

V. Determination of a Lander's Absolute Position

Simultaneous acquisition of Doppler and SBI data, or Doppler and ranging data, from a Mars lander and orbiter enables accurate determination of the lander's absolute position on the surface of Mars. The Doppler data are able to determine the lander components in the plane perpendicular to the Mars spin axis; lander-orbiter SBI, or simultaneous lander-orbiter ranging, constrains the third component of lander position since the orbiter trajectory is tied to the Mars center of mass.

Two different Mars orbiters are considered in this study—a low circular polar orbiter (LCPO), and a half-synchronous circular orbiter (HSCO). The orbital elements for these spacecraft are listed in Table 4. The reference epoch for each orbiter is August 15, 1994 00:00:00 UTC.

Table 5 lists the estimated and considered parameters and a priori uncertainties for the lander-orbiter covariance analysis. The gravity-field uncertainty is based on an analysis of gravity-calibration orbits for Mars Observer early in its mission.¹ The longitude of the lander is fixed, defining a Mars-centered radio-reference frame. Once the z-height and spin radius of this lander are determined, the locations of other landers or rovers in this frame of reference could be established by performing simultaneous Doppler and SBI or Doppler and ranging measurements, as discussed above.

¹ P. B. Esposito, *Mars Observer Navigation Plan: Preliminary*, JPL D-3820 (internal document), Jet Propulsion Laboratory, Pasadena, California, December 16, 1988.

A. Lander Positioning With SBI and Doppler Data From Lander and Orbiter

In the cases studied here, the SBI data are obtained from the lander and a single orbiter, either LCPO or HSCO. As in the lander-rover analysis presented above, an SBI phase bias was estimated in preliminary runs in order to establish that the integer bias could be unambiguously determined.

The lander-orbiter Doppler and SBI data were scheduled for August 16, 1994 from 10:00 through 23:00 UTC. Doppler data from both spacecraft are acquired continuously at Goldstone during this time; SBI data are acquired when two DSN stations can simultaneously view both spacecraft. Figure 6 portrays the Earth-spacecraft visibility during this period. The lander, located at 100-deg longitude and 0-deg latitude, is visible along the Goldstone-Madrid baseline from 10:50 to 15:05, and is visible along the Goldstone-Canberra baseline from 18:40 to 21:35. The LCPO has a period of approximately 130 min, and is occulted by Mars for about 30 min of each revolution. The HSCO is not occulted by Mars during the period of observation.

Because the angular separation of orbiter and lander as viewed from Earth is larger than in the lander-rover case, the SBI data suffer increased degradation from imperfect cancellation of media effects, particularly the solar plasma. Time-tag synchronization uncertainty between the two stations induces an additional error in SBI measurements involving spacecraft whose range rates have a relative drift over the observation [4]. At present, time tags are known to the 0.1- μ sec level. However, nanosecond-level synchronization is feasible and would make this error insignificant. In this study, 8.4-GHz (X-band) lander-orbiter SBI data are weighted at 0.2 mm [3], and 2.3-GHz (S-band) SBI data are weighted at 2.5 mm.

For the lander and the HSCO, Doppler data are weighted as described previously—0.045 mm/sec for 8.4-GHz (X-band) Doppler, 0.45 mm/sec for 2.3-GHz (S-band) Doppler; for the LCPO, Doppler data are *always* weighted at 2.5 mm/sec to reduce sensitivity to uncertainties in the Mars gravity field.

Figure 7 depicts the uncertainties in the lander's z-component and spin radius when Doppler and SBI data are acquired from the lander and orbiter transmitting a 2.3-GHz or 8.4-GHz carrier signal. If the carrier is 2.3-GHz, then simultaneous acquisition of Doppler and SBI from the lander and the LCPO enables determination of the lander's z-component to under 40 m; the error

in the spin radius is about 3 m. If the carrier is 8.4 GHz, the z-component error is below 10 m, while the spin-radius error is at the submeter level. Results of similar quality are obtained when Doppler and SBI are simultaneously acquired from the lander and the HSCO.

B. Lander Positioning With Ranging and Doppler Data From Lander and Orbiter

Simultaneous acquisition of Earth-based ranging and Doppler data also enables very accurate lander-position determination. To study this case, lander and orbiter 8.4-GHz (X-band) Doppler and ranging data were scheduled for August 16 from 10:00 through 23:00 UTC. Data were acquired at a given DSN complex whenever the spacecraft were visible.

Initial covariance analyses indicated that the lander's z-component can be determined to about 130 m if the ranging data have 1-m precision and a 5-m bias. While this level of range data accuracy is representative of the upper limits of the capabilities of the current system, ranging systems with accuracies on the order of tens of centimeters are feasible [15]. For the lander-orbiter cases studied here, the ranging data are assumed to have 30-cm precision, and a 1-m bias due to uncalibrated delays through system hardware.

Figure 8 illustrates the uncertainties in the lander's spin radius and z-height, after 13 hr of 8.4-GHz (X-band) Doppler and precise ranging are simultaneously acquired from the lander and an orbiter. When data are acquired from the lander and either the LCPO or the HSCO, the lander's RSS z-height error is about 30 m. Simultaneous acquisition of Doppler and ranging from the lander and the HSCO enable determination of the lander's spin radius to a few meters. The lander's spin radius can be determined somewhat less accurately if the LCPO is used; the LCPO is more sensitive to gravity mismodeling.

VI. Summary

Radio metric tracking from Earth of multiple spacecraft on the surface of Mars can enable accurate relative and absolute spacecraft-position determination. Simultaneous acquisition of Doppler and ranging enables very accurate relative positioning of Mars landers. Alternatively, SBI combined with Doppler data can enable accurate relative positioning of spacecraft on the surface of Mars with

shorter observation times: Few-meter position determination can be obtained with data arcs as brief as 1 hr. Absolute lander-position determination in Mars-centered coordinates is possible by making differential observations of landers and orbiters. The orbiter provides a tie to the Mars center of mass, allowing the absolute lander z-height coordinate to be determined. Range and Doppler or SBI and Doppler can provide 10- to 30-m accuracy for all components of the absolute lander position.

The range and the SBI data types each have advantages that will determine which technique is most appropriate for a given mission scenario. The combination of Doppler and ranging has the advantage of using only one Earth station at a time, but requires ranging transponders on the Mars surface elements and must be acquired in two-way mode. SBI, on the other hand, requires only a simple one-way carrier signal, and thus could be used with very unsophisticated, low-power surface beacons. The inherently high precision of the SBI observable also allows relative rover-lander positions to be determined extremely rapidly from short data arcs. This could have benefits for quickly obtaining daily updates for lander-rover relative separation.

This study has focused on the ability of Earth-based differential tracking to determine positions for surface elements. In the lander-orbiter case, these same differential observations also serve to benefit determination of the orbiter state. Additional analysis must be performed to quantify the benefits of orbiter-lander relative tracking for Earth-based orbiter navigation.

Over the coming decade, a number of opportunities to demonstrate and apply these differential tracking techniques will arise. The SBI technique is currently being tested through a program of observations of the Magellan and Pioneer Venus Orbiter spacecraft at Venus [4]. In the mid-to-latter part of this decade, the United States' Mars Observer and the former Soviet Union's Mars '94 spacecraft will both be at Mars. Relative tracking of these spacecraft, and possibly of landed elements of the Mars '94 mission, will provide a test-bed for differential tracking which could have potential benefits to both missions. Future robotic and manned Mars missions embodied in the Space Exploration Initiative would ultimately benefit from these techniques. As the infrastructure of Mars tracking and navigation resources grows, in-situ tracking techniques will also come into play; nonetheless, Earth-based tracking will continue to play an important complementary role in the overall Mars navigation challenge.

References

- [1] J. S. Border and W. M. Folkner, "Differential Spacecraft Tracking by Interferometry," *Proceedings of the CNES International Symposium on Space Dynamics*, Nov. 6-10, 1989, Toulouse, France: Cepadues-Editions, pp. 643-654, 1990.
- [2] J. S. Border and R. D. Kahn, "Relative Tracking of Multiple Spacecraft by Interferometry," *Proceedings of the AAS/GSFC International Symposium on Orbital Mechanics and Mission Design*, San Diego, California: Univett, Inc., pp. 441-454, 1989.
- [3] W. M. Folkner and J. S. Border, "Orbiter-Orbiter and Orbiter-Lander Tracking Using Same-Beam Interferometry," *Technical Papers, AIAA/AAS Astrodynamics Conference Part 1* (A90-52957 24-13), Washington, D.C.: American Institute of Aeronautics and Astronautics, pp. 355-363, August 20-22, 1990.
- [4] J. S. Border, W. M. Folkner, R. D. Kahn, and K. S. Zukor, "Precise Tracking of the Magellan and Pioneer Venus Orbiters by Same-Beam Interferometry," paper AAS 91-191 presented at the AAS/AIAA Spaceflight Mechanics Meeting, Houston, Texas, February 1991.
- [5] N. C. Ham, "VLBI System (BLK I) IF-Video Down Conversion Design," *TDA Progress Report 42-79*, vol. July-September, Jet Propulsion Laboratory, Pasadena, California, pp. 172-177, November 15, 1984.
- [6] G. Lindal, H. B. Hotz, D. N. Sweetnam, Z. Shippony, J. P. Brenkle, G. V. Hartsell, and R. T. Spear, "Viking Radio Occultation Measurements of the Atmosphere and Topography of Mars: Data Acquired During 1 Martian Year of Tracking," *Journal of Geophysical Research*, vol. 84, no. B14, pp. 8443-8456, December 30, 1979.
- [7] A. Seiff and D. B. Kirk, "Structure of the Atmosphere of Mars at Mid-Latitudes," *Journal of Geophysical Research*, vol. 82, no. 28, pp. 4364-4378, September 30, 1977.
- [8] A. R. Thompson, J. M. Moran, and G. W. Swenson, Jr., *Interferometry and Synthesis in Radio Astronomy*, New York: John Wiley and Sons, 1986.
- [9] A. Jursa, ed., *Handbook of Geophysics and the Space Environment*, Air Force Geophysics Laboratory, Springfield, Virginia: National Technical Information Service, 1985.
- [10] A. P. Freedman, "Determination of Earth Orientation Using the Global Positioning System," *TDA Progress Report 42-99*, vol. July-September, Jet Propulsion Laboratory, Pasadena, California, November 15, 1989.
- [11] R. D. Kahn and J. S. Border, "Precise Interferometric Tracking of Spacecraft at Low Sun-Earth-Probe Angles," paper 88-0572 presented at the AIAA 26th Aerospace Sciences Meeting, Reno, Nevada, January 11-14, 1988.
- [12] A. P. Mayo, W. T. Blackshear, R. H. Colson, W. H. Michael, Jr., G. M. Kelly, J. P. Brenkle, and T. A. Komarek, "Lander Locations, Mars Physical Ephemeris, and Solar System Parameters: Determination From Viking Lander Tracking Data," *Journal of Geophysical Research*, vol. 82, no. 28, pp. 4297-4303, September 30, 1977.
- [13] A. Konopliv and L. J. Wood, "High-Accuracy Mars Approach Navigation with Radio Metric and Optical Data," paper AIAA-90-2907 presented at the AIAA/AAS Astrodynamics Conference, Portland, Oregon, August 20-22, 1990.

- [14] A. Cazenave and G. Balmino, "Meteorological Effects on the Seasonal Variations of the Rotation of Mars," *Geophysical Research Letters*, vol. 8, no. 3, pp. 245-248, March 1981.
- [15] L. E. Young, "Improved Ranging Systems," taken from *Relativistic Gravitational Experiments in Space*, NASA Conference Publication 3046, pp. 203-205, NASA, Washington, D.C., June 1988.

Table 1. One-second SNR for three different radio frequency configurations.

Carrier frequency, GHz	Case 1 2.3 (S-band)	Case 2 8.4 (X-band)	Case 3 32.5 (Ka-band)
Spacecraft antenna diameter, m	Omni	0.5	0.5
Signal power, W	5	10	10
Ground antenna diameter, m	34	34	34
Earth-Mars distance, AU	2.5	2.5	2.5
Receiver system temperature, K	20	25	80
1-sec SNR _v	3.6	152	361

Table 2. SBI measurement error (in mm) for a Mars lander and rover separated by 100 km. The error is computed for the three different radio frequency configurations listed in Table 1. Integration time is 5 min.

Source of error	Case 1	Case 2	Case 3
System noise	0.664	0.004	<0.001
Phase dispersion	0.361	0.100	0.026
Earth troposphere	<0.001	<0.001	<0.001
Mars troposphere	0.008	0.008	0.008
Earth ionosphere	0.003	<0.001	<0.001
Mars ionosphere	0.050	0.004	<0.001
Station locations/UTPM	<0.001	<0.001	<0.001
Solar plasma (20° Sun-Earth-probe angle)	0.059	0.004	<0.001
Root sum square	0.76	0.10	0.03

Table 3. Estimated and considered parameters for lander-rover covariance analysis.

Estimated parameters	A priori sigma
Lander-Rover relative separation	100 km/component
Right ascension and declination of Mars pole	50 μ rad/angle
Mars UT1-UTC	30 μ rad
Mars, Earth ephemeris corrections	≈ 1 km/component [13]
SBI phase bias (preliminary runs only)	10 ⁶ km
Range bias (Doppler-ranging runs only)	5 m
Considered parameters	A priori sigma
Mars zenith-troposphere delay uncertainty	40 mm
Earth zenith-troposphere delay uncertainty (after calibration)	40 mm
Mars zenith-TEC uncertainty	5 \times 10 ¹⁶ electrons/m ²
Earth zenith-TEC uncertainty (after calibration)	5 \times 10 ¹⁶ electrons/m ²
Earth station locations (including UT and polar motion)	70 mm/component

Table 4. Orbital elements for low circular polar orbiter and half-synchronous circular orbiter.
Reference epoch is August 15, 1994 00:00:00 UTC.

Spacecraft	Semi-major axis, km	Eccentricity	Inclination, deg	Argument of periapsis, deg	Longitude of ascending node, deg	Mean anomaly, deg
LCPO	4150.0	0.0001	91.0	-90.0	0.0	0.0
HSCO	12868.63	0.0001	55.0	-90.0	0.0	0.0

Table 5. Estimated and considered parameters for lander-orbiter covariance analysis.

Estimated parameters	A priori sigma
Lander spin radius and Z-height	100 km per component
Orbiter epoch state position	10^4 km per component
Orbiter epoch state velocity	10 km/sec per component
Right ascension and declination of Mars pole	$50 \mu\text{rad}/\text{angle}$
Mars, Earth ephemeris corrections	≈ 1 km/component [13]
Mars UT1-UTC	$175 \mu\text{rad}$
SBI phase bias (preliminary runs only)	10^6 km
Range bias (Doppler-ranging runs only)	1 m
Considered parameters	A priori sigma
Mars and Earth media	(See Table 3)
Earth station locations	70 mm/component
Solar reflection coefficients	10 percent
Atmospheric drag coefficient	20 percent
Bias acceleration	10^{-12} km/sec ² per component
Mars GM	$3.5 \times 10^{-6} \times \text{GM}$
Mars gravity field	Errors from Mars Observer calibration orbit ²
² Ibid.	

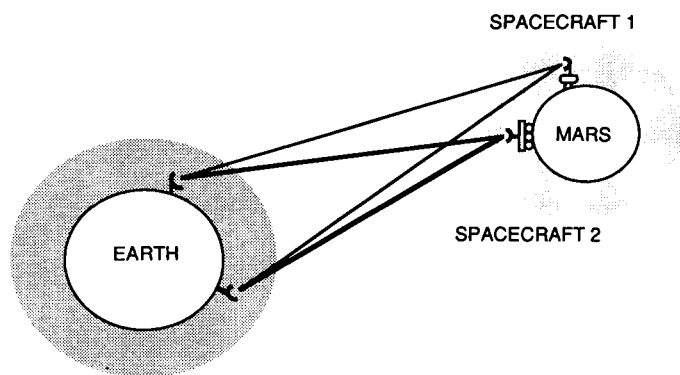


Fig. 1. Same-Beam Interferometry geometry.

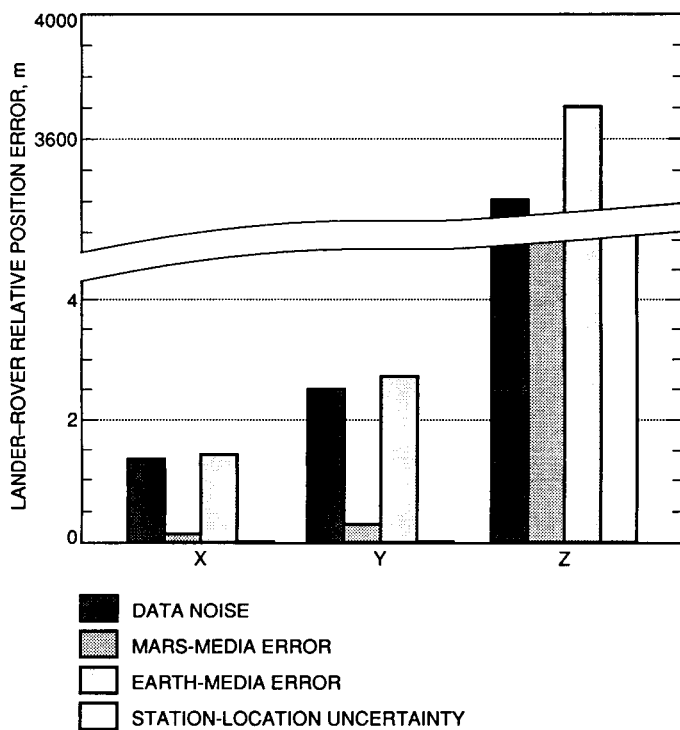


Fig. 2. Lander-rover relative position determination using 3.75 hr of Doppler data (carrier frequency 8.4 GHz).

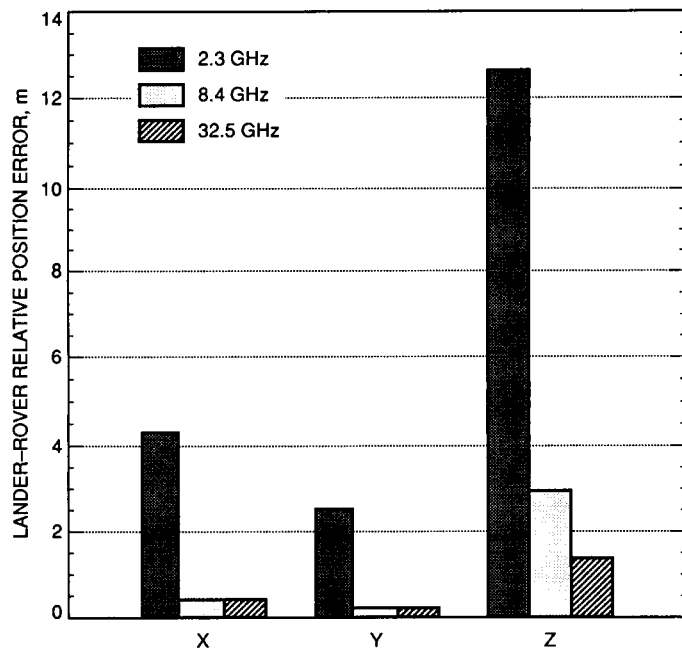


Fig. 3. Lander-rover relative position determination using 3.75 hr of Doppler and Same-Beam Interferometry data. Carrier frequency is 2.3 GHz (S-band), 8.4 GHz (X-band), or 32.5 GHz (Ka-band).

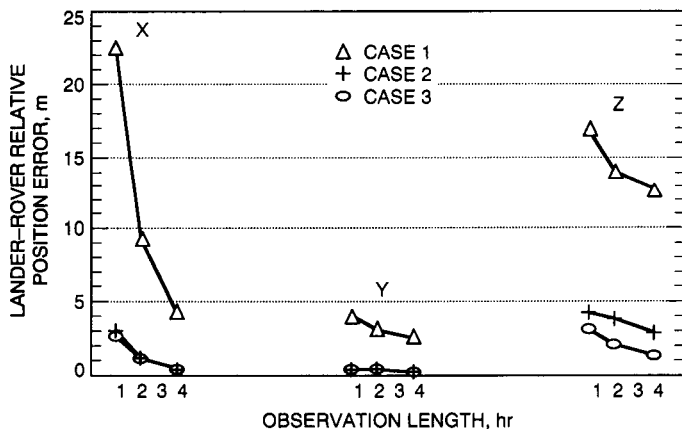


Fig. 4. Doppler and SBI performance as a function of observation length, for each of the cases listed in Table 1. Lander-rover relative position errors are plotted for each component of spacecraft separation after 1 hr, 2 hr, and 3.75 hr.

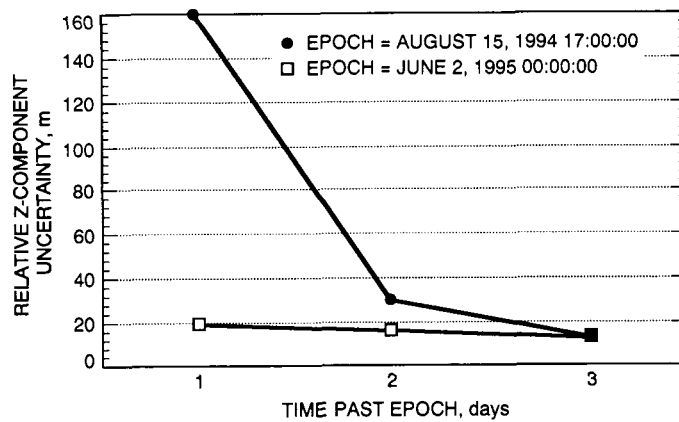


Fig. 5. Uncertainty in z-component of lander-rover relative position when ranging and X-band (8.4-GHz) Doppler are acquired from both spacecraft. On August 15, 1994, the declination of Earth is 3 deg; on June 2, 1995, the declination of Earth is 24 deg.

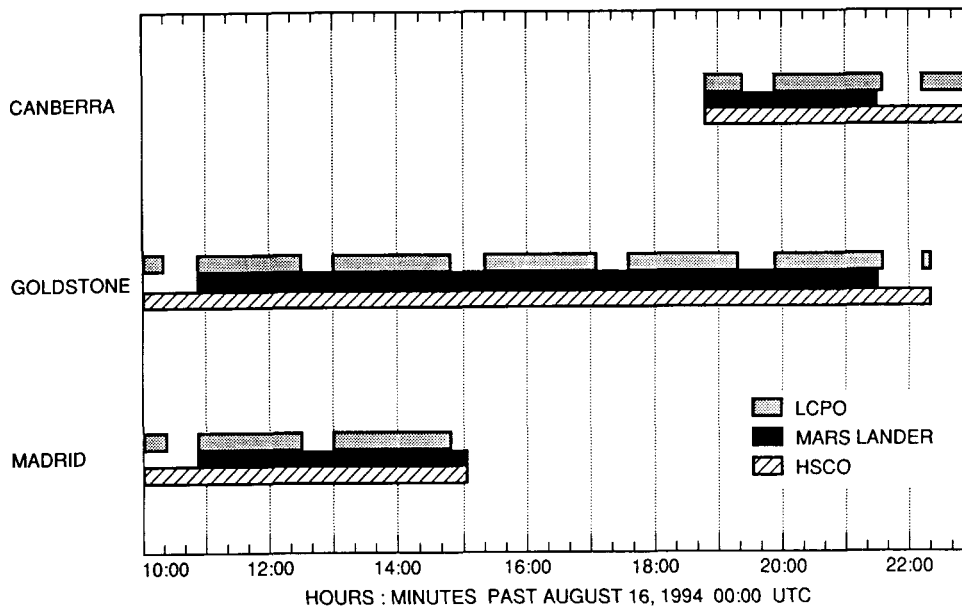


Fig. 6. DSN view periods of low circular polar orbiter, Mars lander, and half-synchronous circular orbiter. The orbital elements are listed in Table 4. The lander is located at 100 deg longitude and 0 deg latitude.

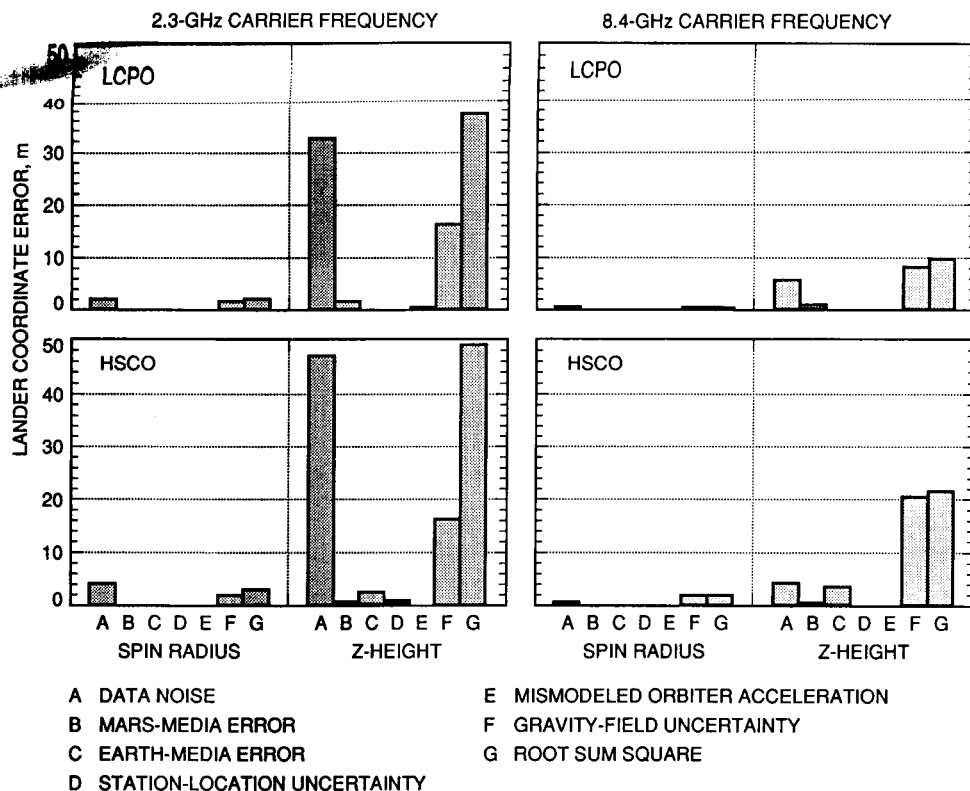


Fig. 7. Error in lander position when 2.3-GHz (S-band) or 8.4-GHz (X-band) Doppler data and SBI data are acquired from the lander and a low circular polar orbiter or a half-synchronous circular orbiter. SBI data are acquired along both the Goldstone-Canberra and the Goldstone-Madrid baselines.

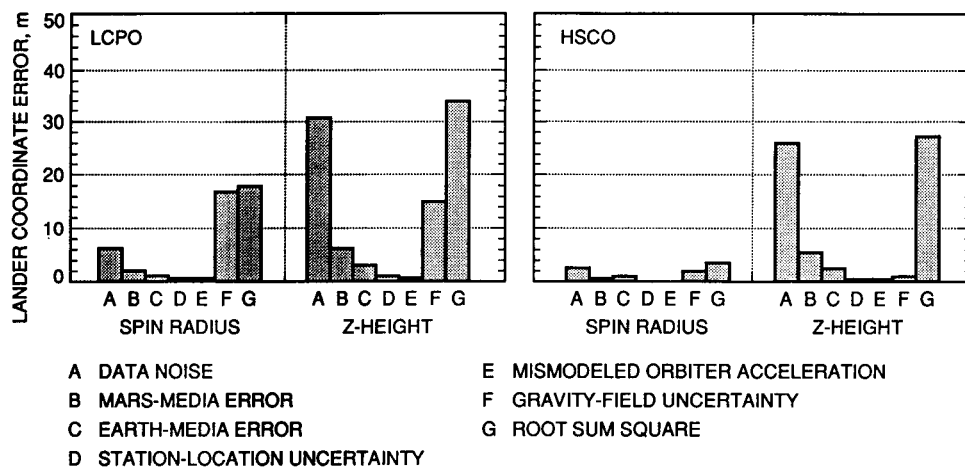


Fig. 8. Error in lander position when 8.4-GHz (X-band) Doppler data and 30-cm ranging are acquired from the lander and a low circular polar orbiter or a half-synchronous circular orbiter.

P-13
N92-24323

J574450

Using Single Buffers and Data Reorganization to Implement a Multi-Megasample Fast Fourier Transform

R. D. Brown

Communications Systems Research Section

Data ordering in large fast-Fourier transforms (FFT's) is both conceptually and implementationally difficult. This article describes a method of visualizing data orderings as vectors of address bits, which enables the engineer to use more efficient data orderings and reduce double-buffer memory designs to single-buffer designs. In particular, this article details the difficulties and algorithmic solutions involved in FFT lengths up to 4 megasamples (Msamples) and sample rates up to 80 MHz. Although the particular solutions mentioned may be directly applicable only to the particular system for which they were intended, the methodology by which these solutions were found could be useful to anyone confronted with similar problems.

I. Introduction

The Search for Extraterrestrial Intelligence (SETI) Program has recently completed the wire-wrap prototype wideband spectrum analyzer (WBSA) [1] and is about to start design of the sky-survey signal processor (SSSP) [2]. Both of these machines are high-speed fast-Fourier transform (FFT) processors followed by special-purpose signal-detection hardware. In the course of building the WBSA, a particular methodology grew for the design of FFT machines (memory boards, in particular). These methods have cut the memory requirements for the proposed SSSP system by about 50 percent.

This article begins with a brief description of the SSSP to familiarize the reader with the system from which examples will be drawn. The backbone of this article is the data-reordering scheme, which is followed by a section outlining the method of replacing double buffers with single buffers. Finally, the design of the SSSP's input buffer (INBUF) board is presented as an illustration of the design techniques.

II. Brief Description of the SSSP System

Figure 1 is a block diagram of one of the eight identical processors that will constitute the SSSP. Each processor of the system accepts complex samples at a rate of 80 MHz and performs an FFT on each group of 4 megasamples (Msamples). Although the processor's internal clock runs at only 40 MHz, the processor can accept data at 80 MHz because it has two input lines, the high-input line and the low-input line. The high data point is the one that was sampled first and held until the low point was sampled. Every 25 nsec (40-MHz rate), two sample points enter the processor, and two frequency values from a previous spectrum leave the processor at the back end of the pipe.

A pipeline configuration was chosen to facilitate real-time processing. To accommodate a 4-megapoint (Mpoint) FFT, the data are configured as a 4096-by-1024 matrix, and the FFT is broken into two orthogonal FFT's, one for the data in each column and one for the data in each row. Each FFT is further broken into radix-4 FFT stages. The two groups of these stages are called super-stages.

Proper FFT implementation demands that the first FFT super-stage be performed on points that have the largest possible sampling interval between them, which indicates the columns of the matrix. Accordingly, the INBUF board transposes the matrix before the first of the two FFT super-stages. The result of the transposition is a 1024-by-4096 matrix, in which the first new row contains the first point of all the old rows, the second new row contains the second points, and so on.

Using the convention that the matrix is always oriented so that FFT's are performed on rows, the first super-stage performs an FFT on each 4-kilopoint (kpoint) row. Then the matrix is transposed by the corner-turn memory (CTM) board to prepare it for the second super-stage. The 1-kpoint FFT's are performed on each of the new rows, resulting in a 4-Mpoint spectrum of frequency bins.

The real adjust (RADJ) board reorganizes (but does not transpose) these frequency values for its own purposes. The unscrambler (UMR) board undoes the shuffling done by the RADJ board and transposes the matrix for a final time, putting the frequencies in sequential order.

III. Relating Data Order to Addressing

Consider the sampled data stream as an array, in the case of the SSSP, a 2^{22} -point array. As such, each data point can be specified by a 22-bit index number. Since the SSSP is a dual-rail system, 2^{21} of the samples come in on the high rail, and the other 2^{21} enter on the low rail. The 22-bit index number can be divided as follows: One bit, known as the hi/lo bit, indicates which rail the data came in on, 0 for high, 1 for low. (As a convention, this bit is always the least-significant bit, lsb.) The remaining 21 bits form a 21-bit binary counter, which increments with each clock cycle. As a design convention, the bits of this data counter are labeled C_0 (lsb) through C_{20} , the most-significant bit (msb).

At each clock cycle in Fig. 2, two data points enter the SSSP, one on the high data rail and one on the low data rail. Each of these data points has its own 22-bit index number, and they must be different. Because they enter the board at the same time, the counter bits (the most-significant 21 bits) are the same. The only difference is the hi/lo bit, which will be a 0 for the data point on the high rail and a 1 for the data point on the low rail. Figure 2 depicts this graphically for three clock cycles in the middle of a spectrum.

At each of the three clock cycles, the value of the time counter is given in both decimal and binary form. Below

the time counter are sections of the high- and low-data channels, one data point per channel per clock cycle. The data-order index of each data point is given in decimal and binary forms inside the data box. Because data in this example are in sequential order, the bits in the time counter are in the same order as they appear in the data-order index. If a different order had been chosen, the bits would have been scrambled, but sequential order is easiest to visualize. In sequential order, the index for data on the high channel will always be exactly double the time counter, and the index for data on the low channel will always be one more than the simultaneous data on the high channel.

Once this 22-bit index number is established for a given data point and the addressing algorithm is known, everything that happens to that data point is also known. The data point is stored in a memory array using the 22-bit number as an address, so it is known where the data are stored. When the data are read out of memory, their channels (high or low) and their positions in the data stream are based on those same 22 bits. In the 4-bit example in Fig. 3, data (in the boxes) enter in sequential order. Above each box is its 4-bit data-order index. The 3-bit time counter always matches the 3 left-most index bits, and the right-most bit indicates the high or low channel. The resultant pattern, A_0 through A_3 , differentiates each data point by identifying its position in the sequential data stream (0–15).

As an example, these points are output at the bottom of Fig. 3 in high-low (HL) order. An N -point data stream in HL order looks like:

HI CHANNEL: 0, 1, 2,
 3, 4, 5, ...

LO CHANNEL: $N/2$, $N/2 + 1$, $N/2 + 2$,
 $N/2 + 3$, $N/2 + 4$, $N/2 + 5$, ...

To organize data points in HL order, the msb of the data-order index moves to the least-significant position. The other bits are all in sequential order, but they are shifted to the left by one bit. Again, the time counter matches with the left-most 3 index bits, but now those bits are A_2 , A_1 , and A_0 . (They were A_3 , A_2 , and A_1 in sequential order.) The right-hand bit that indicates the high or low channel is now A_3 . However, when these bits are written in their correct numerical order, $A_3A_2A_1A_0$,

they are the correct binary representation of each data-order index (the large numbers inside the boxes).

In general, any data-reordering scheme (from sequential order to HL order in the above example) can be viewed as a transformation of the data-index values. In this way, the SSSP data index is viewed as a 22-bit vector and the transformation as a 22-by-22 matrix. The resultant new vector is composed of 22 new data-index values. Each new index value is a function of one or more of the original index values. For simple reordering (as in the above example), each new index value is a function of exactly one old index value, and the transformation is simply a reordering of the bits of the data index. Other, more complex, transformations are also useful.

IV. Addressing Single Buffers and Double Buffers

The SSSP is based on the decision to use single buffers instead of double buffers wherever possible, sacrificing simplicity for the sake of hardware savings, board space, and lower power consumption. In the case of the INBUF board, which will be discussed later, this amounts to 8 Mbytes of memory, saving \$8400 and a maximum of 18 W for each of the 8 copies that will be made. Savings for the CTM board are treble this, a net total of over \$200,000 and 384 W for the 8 copies of the CTM.

Figure 4 shows the way in which the use of single buffers differs from double buffers. In each example, the WRITE line tells which spectrum is being written at any particular time, and the READ line tells which one is being read. (Notice that the first spectrum to be read is unlabeled because it contains no real data and is discarded.) Below the WRITE and READ lines, the addressing lines tell the addressing used for each buffer.

Notice that when using a double buffer, only two addressing orders are required, a read order and a write order. Whenever a spectrum is being read out of one buffer (using the read addressing), another spectrum is being written to the other buffer (using the write addressing). Then the latter buffer is read, while a third spectrum is being written into the former buffer.

For simplicity's sake, it is often convenient to write the incoming data in sequential order and read them out in transformed order; call the transformed order T . If the order of the data can be written as an N -point vector, then the transformation between two orders can be written as an N -by- N matrix. Thus, each data order in Fig. 4 is

labeled as sequential or as a power of the transformation matrix, T , a 5-by-5 matrix. The columns of each matrix correspond to the different address lines of the memory buffer, A to E from top to bottom. The order of these columns mirrors the order of the address bits in the 5-bit data-order index.

It is important to realize that if data are to be read out in sequential order after the transformation, they would have to be written in inverse- T order. This is a very important basic concept: *Data can be written in any order, as long as the transformation between the written order and the read order is T .* For example, if instead of being written sequentially, the data order were transformed so that data were written in W order, they would have to be read out in R order such that $R = T * W$.

This idea is the basis for the single-buffer addressing schemes. Using read-write cycles, in which a given address is read then overwritten, the first spectrum is written sequentially. However, when it is read out later (again using read-write cycles), the second incoming spectrum must be written in the same order that the first is going out. (If new data were written in a different order, some memory locations would be rewritten before they were read.) This means that when Spectrum 0 is read out in transformed order, T , Spectrum 1 is being written in T order.

This is the problem. If the second spectrum is read out in T order, as in the double-buffer example, it would appear in the same order it went in, sequentially. So it must be read out in $T * T$ order. This means that Spectrum 2 will be written in $T * T$ order and read in $T * T * T$ order, Spectrum 3 will be written in $T * T * T$ order, and so on. Since there are only a finite number of ways to reorganize the data stream, the pattern will eventually repeat. Reworded in more rigorous terms, a finite P exists such that $T^P = T^0 = I$, where I is the 5-by-5 identity matrix representing sequential order. However, if P is very large, the cycle will take a long time to repeat, requiring extra counters, and leading to complicated address equations and confused engineers. The trick is to simplify the address patterns as much as possible by minimizing P .

V. Interleaving

This section deals with the practical constraints of the present technology. The first difficulty one is likely to run into is the speed of available memories. The particular memory modules used in this processor have a 100-nsec read/write cycle time. (Faster memories are available but not in sizes that are practical for the amount of data required.) Because the clock period is 25 nsec, new data can

be written only every fourth clock cycle. In the meantime, data must be stored in registers until the start of the next 100-nsec cycle (see Fig. 5). Because 2 new points are ready to be written each clock cycle (25 nsec), 8 points accumulate every 100 nsec. All 8 data points must be written at the same time, requiring 8 memory modules operating in parallel (see Fig. 6).

Likewise, 8 data points are read out during each 100-nsec read/write cycle. Two data points are selected by the multiplexer during each of the next four clock cycles. Meanwhile, another 8 points are being read, and the cycle repeats. The following examples will refer to writing data, but it is to be understood that reading data is an analogous operation that is happening simultaneously.

If the sequential example from Fig. 3 were implemented using Fig. 6, the first point on the high channel would be stored in the H0 memory slice, the second point on the high channel in the H1 memory slice, and so on. Something must designate which memory slice a point will be written to. In this simple example, the answer is obvious. The bottom 2 counter bits can be fed directly to the demultiplexer to specify where to steer the data. (This is a slight oversimplification. In actual implementation, the input registers must include clock enables that depend on the same 2 bits.)

More complex situations demand a more formal approach. The 8 memory slices can be distinguished with a 3-bit number, $M = (M_2, M_1, M_0)$. Since each group of 8 data points is written simultaneously, each point within the group must be written to a different memory slice. Thus, the 8 different data indices must each designate a different memory. It is known that these 8 points all arrived within 4 clocks of each other. That means that all the bits of the indices will be the same except for the 2 least-significant (fastest changing) counter bits and the hi/lo bit. These 3 bits must all be what are called memory-selection bits. The memory-slice number, M , is a function of those 3 bits, and the 8 possible values of the 3 bits must correspond to every possible value of M . To clarify notation, M is the 3-bit number that specifies the memory slice, M_{number} is one of the bits of M , and M_{letter} is a memory selection bit within the data index number. Each M_{number} is a function of one or more M_{letter} bits.

Let the hi/lo bit be named M_a , the least-significant counter bit M_b , and the second counter bit M_c . In the example of Fig. 3, the simplest solution is

$$M_2 = M_a$$

$$M_1 = M_c$$

$$M_0 = M_b$$

Under this scheme, the data from the high channel will fill memory slices 0 to 3 in sequential order, and the data from the low channel will fill memory slices 4 to 7 in order. However, it is perfectly reasonable to use a different ordering. For example, the functions could have been defined

$$M_2 = M_a$$

$$M_1 = M_b$$

$$M_0 = M_b \wedge M_c$$

where \wedge is taken to be the C-language symbol for "exclusive-or."

At time $t = 0$, the high-channel data will be written to memory slice 0 and the low channel to slice 4. At $t = 1$, $M_b = 1$ and $M_c = 0$, so the high-channel data will be written to slice 3 and the low channel to slice 7. At $t = 2$, high-channel data will be written to slice 1 and low data to slice 5. Finally, at $t = 3$, high and low data are written to slices 2 and 6, respectively. The cycle repeats at times $t = 4, 8, 12$, and so on.

VI. INBUF Board

The first (and simplest) memory board in the system is the INBUF board. Data enter sequentially:

HI CHANNEL : 0, 2, 4, 6, 8, ...

LO CHANNEL : 1, 3, 5, 7, 9, ...

This ordering can be represented as 22 address bits, labeled A to V, with V being the least-significant (hi/lo) bit (see Fig. 7). The data can also be looked at as being in a large 4096-by-1024 matrix, without altering the ordering at all. Step 1 is a purely conceptual step. The most-significant 12 bits designate the row number. The least-significant 10 bits specify the column (or position within the row). The row bits are separated from the column bits by a hyphen. As a standard convention, the least-significant bits are always row bits unless otherwise specified.

To fully reap the benefits of doing an FFT, the first stage operates on the samples that are most widely separated in time [3]. Subsequent stages work on data points nearer each other in time, and the last stage operates on sequential samples. In terms of the matrix FFT, the first super-stage should operate on the vector that contains samples $(S_0, S_{1k}, S_{2k}, S_{4k}, S_{6k}, \dots, S_{4Meg-1k})$. This is the first column of the matrix. (The first super-stage operates, of course, on each individual column, not just the first.)

The FFT boards that follow the INBUF operate on groups of 4 kpoints of data at a time. To minimize the amount of memory required within the FFT boards, those 4 kpoints should enter sequentially, so the INBUF board transposes the matrix in Step 2. The old row bits become column bits and vice versa.

The next reorganizational step (Step 3) is dictated by a knowledge of the inner workings of the FFT boards. They have been designed to do a separate FFT on each channel, with no data being shared between the high and low channels. Thus, a given row appears on one channel or the other, not both. This requires that the hi/lo bit be a row bit. The least-significant row bit moves to the hi/lo position.

Step 4 also depends upon the particular architecture of the FFT boards. They contain cascaded stages of radix-4 FFT's, which do 4-point FFT's on groups of the incoming data. The first stage does its work on the following quadruples:

$$\begin{aligned} &(S_0, S_{1Meg}, S_{2Meg}, S_{3Meg}), \\ &(S_1, S_{1Meg+1}, S_{2Meg+1}, S_{3Meg+1}), \\ &(S_2, S_{1Meg+2}, S_{2Meg+2}, S_{3Meg+2}), \text{ etc.} \end{aligned}$$

The order in which these quadruples appear is not important, but the order of the points within each quadruple is. So the two most-significant column bits move to the least-significant positions, creating what is known as high-low radix-4 order.

This completes the required reordering of the INBUF board. If data are written sequentially, A through V (as at the top of Fig. 7), and read using the addressing scheme at the bottom of the figure, M through V , all the necessary operations will be performed in a single step. Unfortunately, if this addressing pattern is implemented using a single buffer, the pattern only repeats every 21 spectra (as

seen in Fig. 8). The worst feature of this particular design is that each address bit appears in a different column each spectrum. Thus, each and every address bit is a function of 21 different data-counter bits and a 5-bit spectrum counter, yielding huge unwieldy logic equations. With a little ingenuity, this can be simplified greatly.

Figure 9 demonstrates the design process of an efficient scheme. Step 1 is the same as before, and Steps 2 and 3 are similar. As noted before, the order in which the quadruples come is irrelevant. In fact, the order in which the rows come is also irrelevant. The only bits that are fixed are the 2 least-significant column bits, A and B . The other column bits are all interchangeable, as are the row bits, so they are all left blank. (Column bits are not interchangeable with row bits because a complete row must go into each channel of the FFT before the next row enters. This ensures that a separate FFT is being done on every single row. Thus, row and column bits are specified as such although they are not specifically identified.)

The goal of this method is to keep the addressing equations as simple as possible; whenever possible, address bits will remain in the same locations. This is possible only in the cases of the 3 least-significant row bits (which become the 3 most-significant column bits) and the least-significant column bit (which becomes the least-significant row bit). These bits are J , K , L , and V , respectively. In Step 5, these bits are assigned to remain in their respective positions.

A and B are the only bits that are constrained to move to particular positions. To minimize the cycle number, N , these two bits must return to their original positions as soon as possible, preferably after the next cycle. Thus, it is desirable that whichever bits lie in the second and third positions from the right in one spectral addressing will move to the two most-significant positions one spectrum later. By extension, T and U , which were in the second and third positions before the transformation, are required to be in the most-significant positions after the transformation. See Step 6.

By Step 7, all the constraints of the INBUF board have been satisfied, so a column order that is beneficial to the FFT boards is used. (This particular order allows the four FFT boards that comprise the two stages of the FFT to be built as identical copies of each other and minimizes the amount of memory needed to reorganize the data between each internal stage.) Once this order has been determined, the rest of the design follows from matching pairs of address bits in the same manner that A and T were matched and B and U were matched. This ordering scheme repeats after two spectra, which can easily be verified.

Now the design of the transformation is complete, but the implementational design remains. The implementational phase occurs when the 22 data-index bits, originally dubbed A through V , are designated as memory-selection bits or addressing bits. Because the design requires an eight-way memory interleave, three of the bits will be reserved for memory selection. The memories are each 512 kpoints deep, requiring 19 addressing bits (although some of these will also be used for memory selection).

Figures 10 and 11 describe the process of renaming the data-index bits. A bit chosen as an addressing bit is renamed A_{number} , and a memory-selection bit is renamed M_{letter} . These are just new names for the bits that were previously designated A through V . The old names were only placeholders; the new names also give information about how the transformation will be implemented.

There is a lot of freedom here, but the most straightforward choice is to assign the 3 memory-selection bits in the right-hand positions and the 19 address bits in order to the left (see Fig. 10). This works fine for spectrum 0. However, when the address bits have been transformed for spectrum 1, only one memory-selection bit remains in the three right-hand spots. This means that when a group of 8 data points comes in, the board will attempt to write them all to only two memory slices.

This problem is alleviated by stipulating that A and B must also be memory-selection bits as well as address bits (see Fig. 11). M_0 is now a function of both A and T , "exclusive-or"-ing bit A with bit T . This ensures that out of each group of 8 data points, half will be directed to even memory slices and half to odd memory slices. Likewise, B is "exclusive-or"-ed with U to determine M_1 . Now each group of four clocks (100 nsec) accesses all eight memory slices. Note: Only groups that start on a 100-nsec boundary ($C_1 = C_0 = \text{hi/lo} = 0$) need to be considered.

Now the design of the INBUF is complete. The Boolean equations for each address bit can be read directly off of Fig. 11. For instance, A_0 is equivalent to C_2 during spectrum 0, and it is equivalent to C_{17} during spectrum 1:

$$A_0 = (C_2 * S') + (C_{17} * S)$$

where S is the 1-bit spectrum counter, C_{number} is a bit selected from the 21-bit data counter, ' is logical "not," + is logical "or," and * is logical "and." Similar equations can be found for A_1 to A_{16} .

Notice that two address bits occur in the least-significant three places during spectrum 1. This causes

some conceptual difficulties and requires different address lines to different slices of memory. But remember that A_{18} is defined as being equivalent to M_e , which in turn is equivalent to $M_0 \wedge M_e$, where \wedge is "exclusive-or." Thus:

$$A_{18} = C_{20}$$

for memory slices 0, 2, 4, and 6, and

$$A_{18} = (C_{20} * S') + (C_{20'} * S)$$

for memory slices 1, 3, 5, and 7. Equations for A_{17} can be similarly derived.

VII. Different Transformations

All previously mentioned transformations reorganized the address bits but leave the bits unchanged. Sometimes, the designer may wish to read out the data using address bits that are functions of, but not identical to, the original address bits. The simplest example is reverse order. If data are written sequentially, using address bits A , B , and C , they can be read out in reverse order by just inverting the address bits, as shown below:

Counter (A, B, C)	000	001	010	011	...
Data in	0	1	2	3	...
Reverse counter (A', B', C')	111	110	101	100	...
Data out	7	6	5	4	...

Another useful transformation is the downcounter transformation, which is easily implemented by substituting a downcounter for the regular counter. The bits of the downcounter are indicated as \tilde{A} , \tilde{B} , \tilde{C} .

Counter (A, B, C)	000	001	010	011	...
Data in	0	1	2	3	...
Downcounter ($\tilde{A}, \tilde{B}, \tilde{C}$)	000	111	110	101	...
Data out	0	7	6	5	...

These two operations have one thing in common; they are reversible. That means that after the address bits are transformed, no information is lost; the original address bits can still be recovered. Reversibility is important because it prevents memory slices from attempting to write

multiple data points to the same spot. There are many reversible operations, such as "exclusive-or"-ing two bits (as in Fig. 11).

VIII. Limitations

The use of single buffers is limited by two considerations. First, the interleaving required to implement read/write cycles at the proper speed may require as much memory as a double buffer. In such a case, no savings would be realized by using a single buffer.

The other consideration is the complexity of the addressing, which may rise to a point of impracticality.

IX. Conclusion

Visualizing a data stream as a vector of addressing bits allows the designer to treat data reorganization as a matrix transformation on that vector. This allows the designer to easily manipulate many transformations at once and find data orderings that are beneficial to other boards in the system. It also allows the designer to find address patterns that make single-buffer memory banks possible (as opposed to double-buffer banks), further reducing the necessary hardware.

Once an address bit transformation has been chosen, the implementation can be easily designed using methods outlined in this article.

References

- [1] M. Quirk, M. Garyantes, H. Wilck, and M. Grimm, "A Wide-Band High-Resolution Spectrum Analyzer," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, no. 12, pp. 1854-1861, December 1988.
- [2] M. Garyantes, M. Grimm, and G. Zimmerman, "A Wideband High Resolution Digital Spectrum Analyzer for the Search for Extraterrestrial Intelligence," in preparation for the 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing.
- [3] L. R. Rabiner and B. Gold, *Theory and Application of Digital Signal Processing*, Englewood Cliffs, New Jersey: Prentice-Hall, pp. 356-371, 1975.

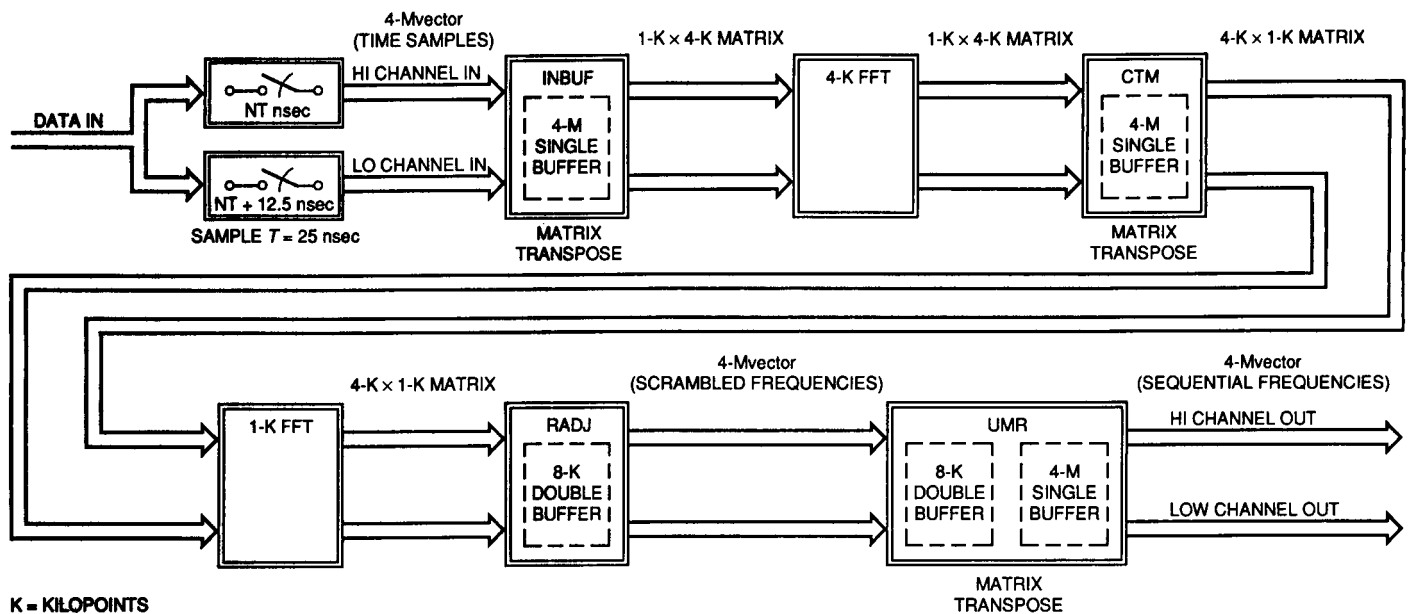


Fig. 1. SSSP.

	TIME = 2565 = 000000000101000000101	TIME = 2566 = 000000000101000000110	TIME = 2567 = 000000000101000000111
HIGH CHANNEL	INDEX = 5130 = 0000000001010000001010	INDEX = 5132 = 0000000001010000001100	INDEX = 5134 = 0000000001010000001110
LOW CHANNEL	INDEX = 5131 = 0000000001010000001011	INDEX = 5133 = 0000000001010000001101	INDEX = 5135 = 0000000001010000001111

Fig. 2. Relating the time counter to the index number.

INPUT - SEQUENTIAL ORDER ($A_3 A_2 A_1 A_0$)							
	$T = 000$	$T = 001$	$T = 010$	$T = 011$	$T = 100$	$T = 101$	$T = 110$
HIGH CHANNEL ($H/L = 0$) $A_3 A_2 A_1 A_0 =$	0000	0010	0100	0110	1000	1010	1100
	0	2	4	6	8	10	12
LOW CHANNEL ($H/L = 1$) $A_3 A_2 A_1 A_0 =$	0001	0011	0101	0111	1001	1011	1101
	1	3	5	7	9	11	13
OUTPUT - HL ORDER ($A_2 A_1 A_0 A_3$)							
	$T = 000$	$T = 001$	$T = 010$	$T = 011$	$T = 100$	$T = 101$	$T = 110$
HIGH CHANNEL ($H/L = 0$) $A_2 A_1 A_0 A_3 =$	0000	0010	0100	0110	1000	1010	1100
	0	1	2	3	4	5	6
LOW CHANNEL ($H/L = 1$) $A_2 A_1 A_0 A_3 =$	0001	0011	0101	0111	1001	1011	1101
	8	9	10	11	12	13	14

Fig. 3. Relating input order to a binary counter.

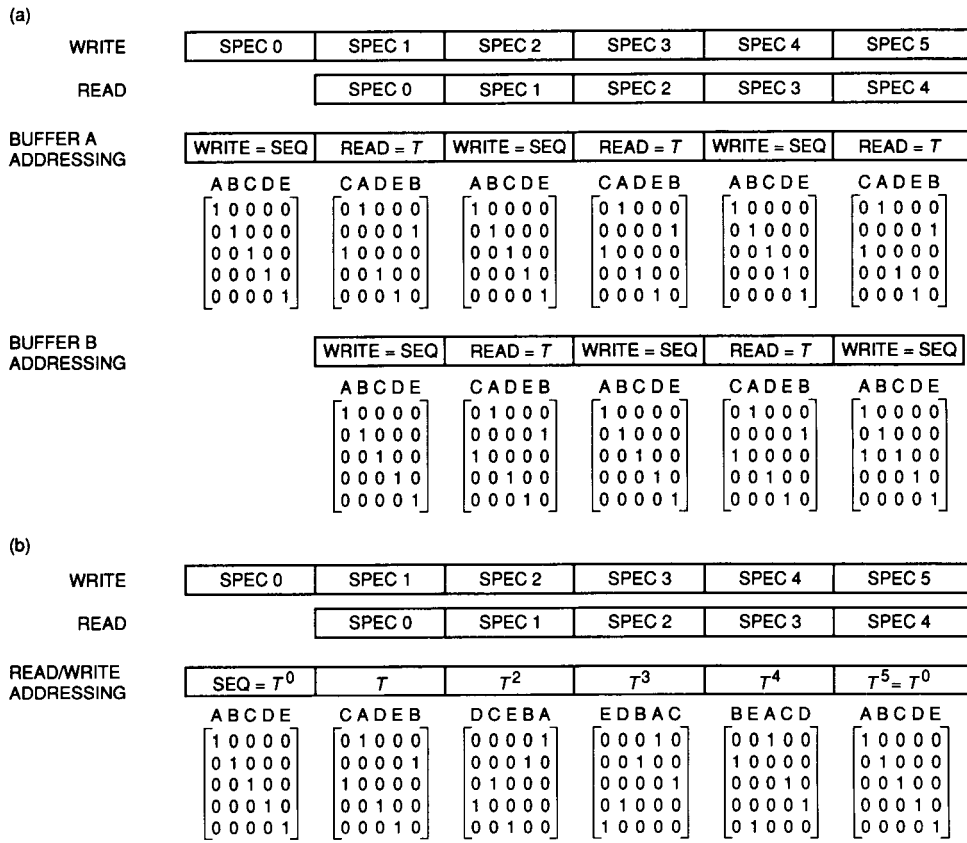


Fig. 4. How to address memories: (a) using a double buffer and (b) using a single buffer.

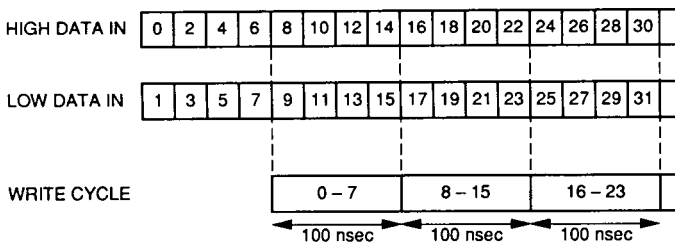


Fig. 5. Timing of eight-way interleave.

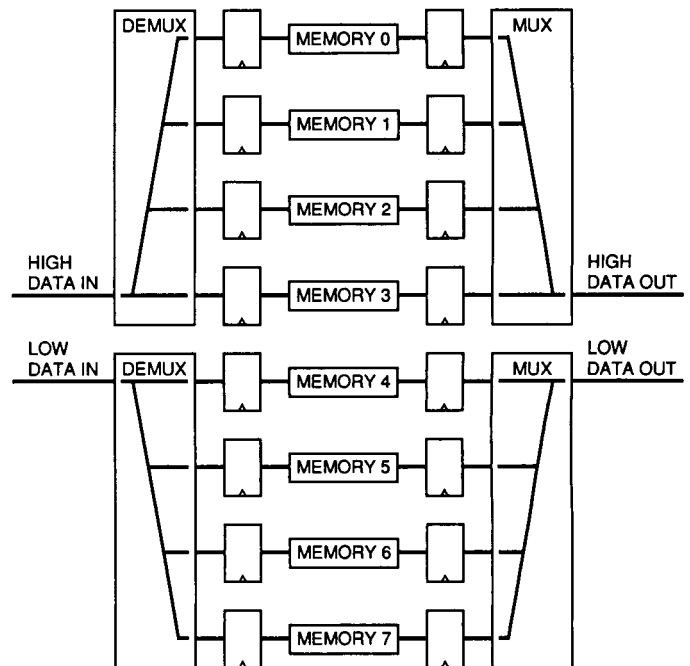


Fig. 6. Memory buffer divided into eight slices.

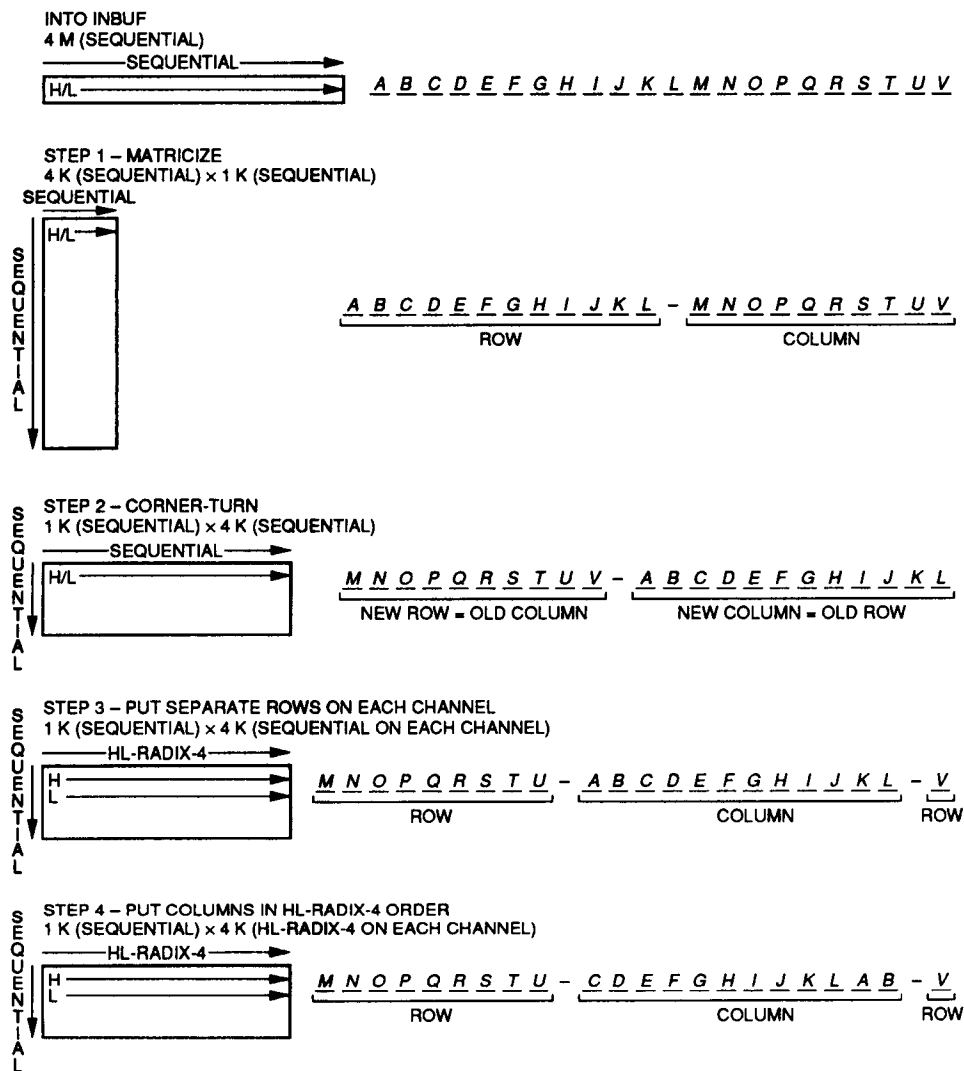


Fig. 7. INBUF data reordering.

	C ₂₀	C ₁₉	C ₁₈	C ₁₇	C ₁₆	C ₁₅	C ₁₄	C ₁₃	C ₁₂	C ₁₁	C ₁₀	C ₉	C ₈	C ₇	C ₆	C ₅	C ₄	C ₃	C ₂	C ₁	C ₀	HI/LO
SPECTRUM = 0	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
SPECTRUM = 1	M	N	O	P	Q	R	S	T	U	C	D	E	F	G	H	I	J	K	L	A	B	V
SPECTRUM = 2	F	G	H	I	J	K	L	A	B	O	P	Q	R	S	T	U	C	D	E	M	N	V
SPECTRUM = 3	R	S	T	U	C	D	E	M	N	H	I	J	K	L	A	B	O	P	Q	F	G	V
SPECTRUM = 4	K	L	A	B	O	P	Q	F	G	T	U	C	D	E	M	N	H	I	J	R	S	V
SPECTRUM = 5	D	E	M	N	H	I	J	R	S	A	B	O	P	Q	F	G	T	U	C	K	L	V
SPECTRUM = 6	P	Q	F	G	T	U	C	K	L	M	N	H	I	J	R	S	A	B	O	D	E	V
SPECTRUM = 7	I	J	R	S	A	B	O	D	E	F	G	T	U	C	K	L	M	N	H	P	Q	V
SPECTRUM = 8	U	C	K	L	M	N	H	P	Q	R	S	A	B	O	D	E	F	G	T	I	J	V
SPECTRUM = 9	B	O	D	E	F	G	T	I	J	K	L	M	N	H	P	Q	R	S	A	U	C	V
SPECTRUM = 10	N	H	P	Q	R	S	A	U	C	D	E	F	G	T	I	J	K	L	M	B	O	V
SPECTRUM = 11	G	T	I	J	K	L	M	B	O	P	Q	R	S	A	U	C	D	E	F	N	H	V
SPECTRUM = 12	G	A	U	C	D	E	F	N	H	I	J	K	L	M	B	O	P	Q	R	G	T	V
SPECTRUM = 13	L	M	B	O	P	Q	R	G	T	U	C	D	E	F	N	H	I	J	K	S	A	V
SPECTRUM = 14	E	F	N	H	I	J	K	S	A	B	O	P	Q	R	G	T	U	C	D	L	M	V
SPECTRUM = 15	Q	R	G	T	U	C	D	L	M	N	H	I	J	K	S	A	B	O	P	E	F	V
SPECTRUM = 16	J	K	S	A	B	O	P	E	F	G	T	U	C	D	L	M	N	H	I	Q	R	V
SPECTRUM = 17	C	D	L	M	N	H	I	Q	R	S	A	B	O	P	E	F	G	T	U	J	K	V
SPECTRUM = 18	O	P	E	F	G	T	U	J	K	L	M	N	H	I	Q	R	S	A	B	C	D	V
SPECTRUM = 19	H	I	Q	R	S	A	B	C	D	E	F	G	T	U	J	K	L	M	N	O	P	V
SPECTRUM = 20	T	U	J	K	L	M	N	O	P	Q	R	S	A	B	C	D	E	F	G	H	I	V
SPECTRUM = 0	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V

Fig. 8. INBUF transformation sequence.

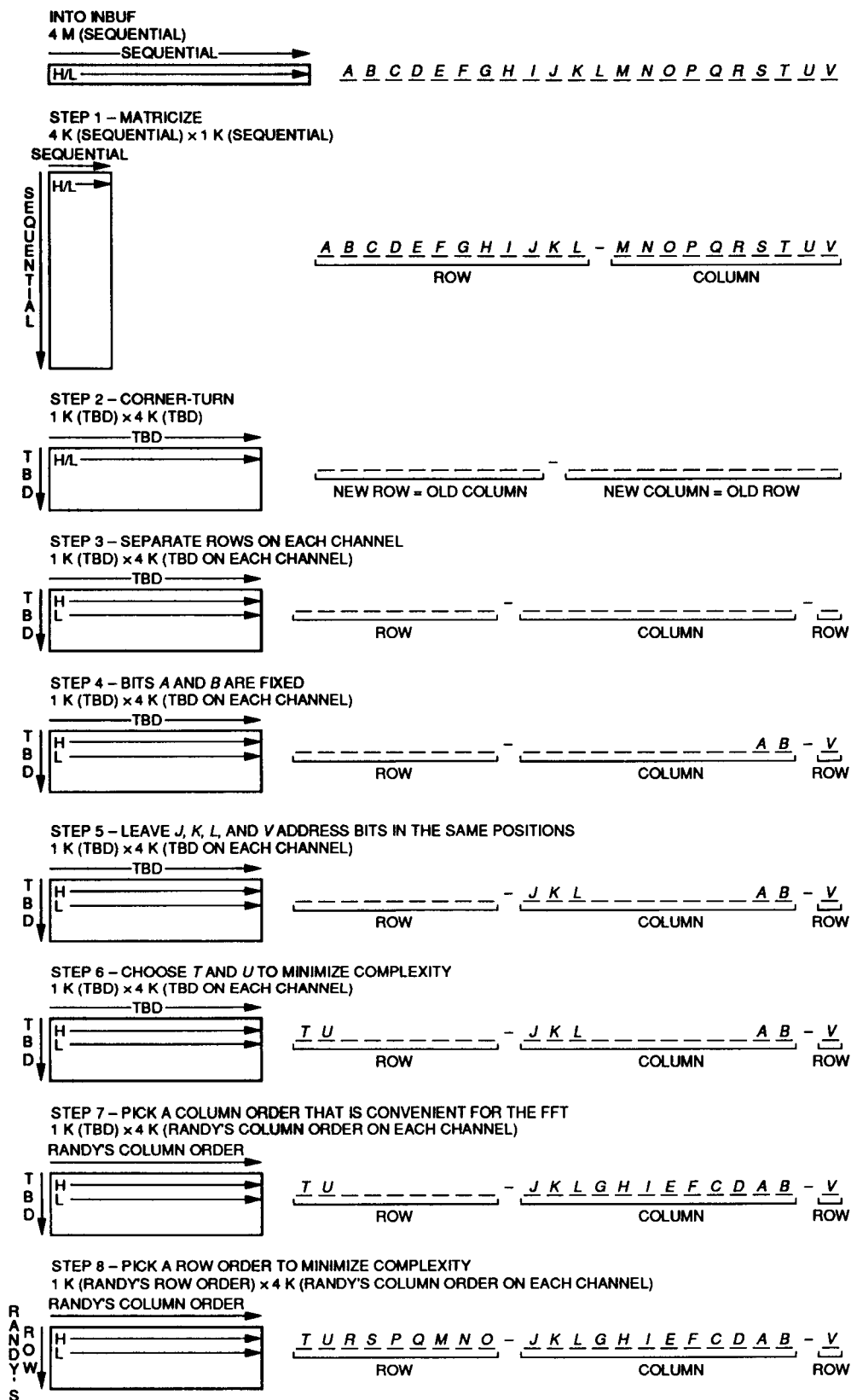


Fig. 9. Revised INBUF data reordering.

	C_{20}	C_{19}	C_{18}	C_{17}	C_{16}	C_{15}	C_{14}	C_{13}	C_{12}	C_{11}	C_{10}	C_9	C_8	C_7	C_6	C_5	C_4	C_3	C_2	C_1	C_0	H/I/O
SPECTRUM = 0	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
	A_{18}	A_{17}	A_{16}	A_{15}	A_{14}	A_{13}	A_{12}	A_{11}	A_{10}	A_9	A_8	A_7	A_6	A_5	A_4	A_3	A_2	A_1	A_0	M_c	M_b	M_a
	$M_2 = M_a$																					
	$M_1 = M_b$																					
	$M_0 = M_c$																					
SPECTRUM = 1	T	U	R	S	P	Q	M	N	O	J	K	L	G	H	I	E	F	C	D	A	B	V
	M_c	M_b	A_1	A_0	A_3	A_2	A_6	A_5	A_4	A_9	A_8	A_7	A_{12}	A_{11}	A_{10}	A_{14}	A_{13}	A_{16}	A_{15}	A_{18}	A_{17}	M_a
	$M_2 = M_a$																					
	$M_1 = M_b$																					
	$M_0 = M_c$																					

Fig. 10. Revised INBUF transformation sequence.

	C_{20}	C_{19}	C_{18}	C_{17}	C_{16}	C_{15}	C_{14}	C_{13}	C_{12}	C_{11}	C_{10}	C_9	C_8	C_7	C_6	C_5	C_4	C_3	C_2	C_1	C_0	H/I/O	
SPECTRUM = 0	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	
	A_{18}	A_{17}	A_{16}	A_{15}	A_{14}	A_{13}	A_{12}	A_{11}	A_{10}	A_9	A_8	A_7	A_6	A_5	A_4	A_3	A_2	A_1	A_0				
	M_e	M_d																			M_c	M_b	M_a
SPECTRUM = 1	T	U	R	S	P	Q	M	N	O	J	K	L	G	H	I	E	F	C	D	A	B	V	
			A_1	A_0	A_3	A_2	A_6	A_5	A_4	A_9	A_8	A_7	A_{12}	A_{11}	A_{10}	A_{14}	A_{13}	A_{16}	A_{15}	A_{18}	A_{17}		
	M_c	M_b																			M_e	M_d	M_a
	</																						

Fig. 11. Final INBUF transformation sequence.

Author Index,¹ 1991

The Telecommunications and Data Acquisition Progress Report

42-105, January–March, May 15, 1991

42-106, April–June, August 15, 1991

42-107, July–September, November 15, 1991

42-108, October–December, February 15, 1992

Abdel-Ghaffar, K.

42-107 Some Partial-Unit-Memory Convolutional Codes, pp. 57–72.

R. J. McEliece and G. Solomon

Abdel-Hamid, T.

42-108 Software-Engineering Process Simulation Model (SEPS), pp. 165–185.

See Lin, C. Y.

Alvarez, L. S.

42-106 Efficiency Calibration of the DSS 13 34-Meter Diameter Beam Waveguide Antenna at 8.45 and 32 GHz, pp. 283–297.

See Slobin, S. D.

42-106 Initial Pointing Calibrations for the DSS 13 34-Meter Beam-Waveguide Antenna, pp. 188–204.

Arbel, D.

42-106 Design and Testing of an Active Quenching Circuit for an Avalanche Photodiode Photon Detector, pp. 102–110.

J. A. Schwartz

Arnold, S.

42-108 Node Synchronization Schemes for the Big Viterbi Decoder, pp. 196–200.

See Cheung, K.-M.

Asmar, S.

42-105 Radio Science Ground Data System for the Voyager–Neptune Encounter—Part I, pp. 109–127.

See Kursinski, E. R.

Bartos, R. D.

42-108 Dynamic Modeling of the Servovalves Incorporated in the Servo Hydraulic System of the 70-Meter DSN Antennas, pp. 222–234.

¹ In the case of joint authorship, the reader is referred to the citation under the first author, where all the authors of the article are listed.

Bathker, D. A.

- 42-105 Gaussian Beam and Physical Optics Iteration Technique for Wideband Beam Waveguide Feed Design, pp. 128-135.

See Veruttipong, W.

Bautista, J. J.

- 42-106 Performance of a High T_c Superconducting Ultra-Low Loss Microwave Stripline Filter, pp. 60-65.

G. Ortiz, C. Zahopoulos, S. Sridhar, and M. Lanagan

Bevan, R.

- 42-106 The Advanced Receiver II Telemetry Test Results at Goldstone, pp. 119-131.

See Sadr, R.

Bienkiewicz, B.

- 42-108 Pointing-Error Simulations of the DSS-13 Antenna Due to Wind Disturbances, pp. 109-134.

See Gawronski, W.

Blewitt, G.

- 42-108 A Demonstration of Centimeter-Level Monitoring of Polar Motion With the Global Positioning System, pp. 1-9.

See Lindqwister, U. J.

Briggs, H. C.

- 42-107 Mark IVA Antenna Control System Data Handling Architecture Study, pp. 96-103.

D. B. Eldred

Britcliffe, M. J.

- 42-106 Efficiency Calibration of the DSS 13 34-Meter Diameter Beam Waveguide Antenna at 8.45 and 32 GHz, pp. 283-297.

See Slobin, S. D.

Brockman, M. H.

- 42-106 A Carrier-Arraying Demonstration at Goldstone for Receiving Pioneer 11 Signals, pp. 307-334.

See Pham, T. T.

Brown, R. D.

- 42-108 Using Single Buffers and Data Reorganization to Implement a Multi-Megasample Fast Fourier Transform, pp. 294-306.

Buchanan, H. R.

- 42-106 Modeling Preparation Costs for Space Missions by Using Major Cost Drivers, pp. 404-415.

See Sherif, J. S.

Charny, B.

- 42-107 A 640-MHz 32-Megachannel Real-Time Polyphase-FFT Spectrum Analyzer, pp. 132-140.

See Zimmerman, G. A.

Chen, J. C.

- 42-105 Gaussian Beam and Physical Optics Iteration Technique for Wideband Beam Waveguide Feed Design, pp. 128-135.

See Veruttipong, W.

- 42-105 X-/Ka-Band Dichroic Plate Design and Grating Lobe Study, pp. 21-30.

- 42-106 DSS 13 Phase II Pedestal Room Microwave Layout, pp. 298-306.

See Cwik, T.

Cheung, K.-M.

- 42-105 Integer Cosine Transform for Image Compression, pp. 45-53.

F. Pollara and M. Shahshahani

- 42-106 Frame Error Rate of the NASA Concatenated Coding System, pp. 183-187.

See Swanson, L.

- 42-107 A Comparison of the Fractal and JPEG Algorithms, pp. 21-26.
See Shahshahani, M.
- 42-108 Node Synchronization Schemes for the Big Viterbi Decoder, pp. 186-200.
L. Swanson and S. Arnold
- Clauss, R.**
- 42-108 DSN G/T_{op} and Telecommunications System Performance, pp. 271-278.
See Stelzried, C.
- Clements, P. A.**
- 42-108 An Assembly to Provide a Time Reference for the JPL Network Operations Control Center, Real-Time pp. 214-221.
J. G. Lopez and M. J. Kelley
- Collins, O.**
- 42-107 A Minimalist Approach to Receiver Architecture, pp. 27-35.
- 42-107 Determinate-State Convolutional Codes, pp. 36-56.
M. Hizlan
- Cormier, R.**
- 42-108 250-kW CW Klystron Amplifier for Planetary Radar, pp. 201-213.
A. Mizuhara
- Cowles, K.**
- 42-105 Site Comparison for Optical Visibility Statistics in Southern California, pp. 31-40.
- 42-108 Alignment of the Atmospheric Visibility Monitoring Telescope, pp. 79-83.
D. Erickson
- Cwik, T.**
- 42-106 DSS 13 Phase II Pedestal Room Microwave Layout, pp. 298-306.
J. C. Chen
- Dick, G. J.**
- 42-107 A Receiver Design for the Superconducting Cavity-Maser Oscillator, pp. 1-11.
See Wang, R. T.
- 42-108 A High-Performance Hg^+ Trapped Ion Frequency Standard, pp. 10-18.
See Prestage, J. D.
- Edwards, C. D.**
- 42-108 Position Determination of a Lander and Rover at Mars With Earth-Based Differential Tracking, pp. 279-293.
See Kahn, R. D.
- Eldred, D. B.**
- 42-107 Mark IVA Antenna Control System Data Handling Architecture Study, pp. 96-103.
See Briggs, H. C.
- Erickson, D.**
- 42-108 Alignment of the Atmospheric Visibility Monitoring Telescope, pp. 79-83.
See Cowles, K.
- Estefan, J. A.**
- 42-106 Precise Orbit Determination of High-Earth Elliptical Orbiters Using Differenced Doppler and Range Measurements, pp. 1-22.
- Flanagan, M. J.**
- 42-107 The Behavior of Quantization Spectra as a Function of Signal-to-Noise Ratio, pp. 155-168.

Folkner, W. M.

- 42-108 Position Determination of a Lander and Rover at Mars With Earth-Based Differential Tracking, pp. 279-293.

See Kahn, R. D.

Franco, M. M.

- 42-105 A Portable X-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation—Part II: Tests on the Antenna, pp. 54-68.

See Otoshi, T. Y.

- 42-106 A Portable Ka-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation—Part I: Design and Ground Tests, pp. 249-265.

See Otoshi, T. Y.

- 42-106 A Portable Ka-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation—Part II: Tests on the Antenna, pp. 266-282.

See Otoshi, T. Y.

- 42-106 Efficiency Calibration of the DSS 13 34-Meter Diameter Beam Waveguide Antenna at 8.45 and 32 GHz, pp. 283-297.

See Slobin, S. D.

- 42-107 A Portable Ku-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation, pp. 73-80.

See Otoshi, T. Y.

- 42-108 The Electrical Conductivities of the DSS-13 Beam-Waveguide Antenna Shroud Material and Other Antenna Reflector Surface Materials, pp. 154-164.

See Otoshi, T. Y.

Freedman, A. P.

- 42-105 Combining GPS and VLBI Earth-Rotation Data for Improved Universal Time, pp. 1-12.

- 42-108 A Demonstration of Centimeter-Level Monitoring of Polar Motion With the Global Positioning System, pp. 1-12.

See Lindqwister, U. J.

Garyantes, M. F.

- 42-107 A 640-MHz 32-Megachannel Real-Time Polyphase-FFT Spectrum Analyzer, pp. 132-140.

See Zimmerman, G. A.

Gawronski, W.

- 42-105 Elevation Control System Model for the DSS 13 Antenna, pp. 83-108.

J. A. Mellstrom

- 42-106 Modeling and Simulations of the DSS 13 Antenna Control System, pp. 205-248.

J. A. Mellstrom

- 42-108 Pointing-Error Simulations of the DSS-13 Antenna Due to Wind Disturbances, pp. 109-134.

B. Bienkiewicz and R. E. Hill

Goldstein, R. M.

- 42-105 Orbital Debris Radar Instrumentation, pp. 146-148.

L. W. Randolph

Gray, D. L.

- 42-106 Δ VLBI Data Performance in the Galileo Spacecraft Earth Flyby of December 1990, pp. 335-352.

Grimm, M. J.

- 42-107 A 640-MHz 32-Megachannel Real-Time Polyphase-FFT Spectrum Analyzer, pp. 132-140.

See Zimmerman, G. A.

Gulkis, S.

- 42-107 Polyphase-Discrete Fourier Transform Spectrum Analysis for the Search for Extraterrestrial Intelligence Sky Survey, pp. 141-154.

See Zimmerman, G. A.

Hill, R. E.

- 42-108 Pointing-Error Simulations of the DSS-13 Antenna Due to Wind Disturbances, pp. 109-134.

See Gawronski, W.

Hinedi, S.

- 42-106 The Advanced Receiver II Telemetry Test Results at Goldstone, pp. 119-131.

See Sadr, R.

- 42-106 Symbol Lock Detection in the ARX II and Block V Receivers, pp. 132-154.

See Shihabi, M.

Hizlan, M.

- 42-107 Determinate-State Convolutional Codes, pp. 36-56.

See Collins, O.

Issa, T. N.

- 42-106 Stochastic Availability Analysis of Operational Data Systems in the Deep Space Network, pp. 394-403.

Kahn, R.

- 42-108 Position Determination of a Lander and Rover at Mars With Earth-Based Differential Tracking, pp. 279-293.

W. M. Folkner, C. D. Edwards, and
A. Vijayaraghavan

Kelley, M. J.

- 42-108 An Assembly to Provide a Time Reference for the JPL Network Operations Control Center, Real-Time, pp. 214-221.

See Clements, P. A.

Kent, S. S.

- 42-106 A Carrier-Arraying Demonstration at Goldstone for Receiving Pioneer 11 Signals, pp. 307-334.

See Pham, T. T.

Kursinski, E. R.

- 42-105 Radio Science Ground Data System for the Voyager-Neptune Encounter—Part I, pp. 109-127.

S. Asmar

Lam, B.

- 42-106 Design and Status of the RF-Digitizer Integrated Circuit, pp. 155-159.

See Rayhrer, B.

Lanagan, M.

- 42-106 Performance of a High T_c Superconducting Ultra-Low Loss Microwave Stripline Filter, pp. 60-65.

See Bautista, J. J.

Lee, P.

- 42-106 Wideband Waveguide Polarizer Development for SETI, pp. 384-393.

P. H. Stanton

Levy, R.

- 42-105 Computation of Vibration Mode Elastic-Rigid and Effective Weight Coefficients From Finite-Element Computer Program Output, pp. 13-20.

- 42-106 Algorithms for Structural Natural-Frequency Design, pp. 23-29.

Lin, C. Y.

- 42-108 Software-Engineering Process Simulation Model (SEPS), pp. 165-185.

T. Abdel-Hamid and J. S. Sherif

Lindqwister, U. J.

- 42-108 A Demonstration of Centimeter-Level Monitoring of Polar Motion With the Global Positioning System, pp. 1-9.

A. P. Freedman and G. Blewitt

Lopez, J. G.

- 42-108 An Assembly to Provide a Time Reference for the JPL Network Operations Control Center, Real-Time, pp. 214-221.

See Clements, P. A.

Maleki, L.

- 42-108 A High-Performance Hg⁺ Trapped Ion Frequency Standard, pp. 10-18.

See Prestage, J. D.

Matousek, S. E.

- 42-107 Trajectory and Navigation System Design for Robotic and Piloted Missions to Mars, pp. 113-131.

See Thurman, S. W.

McEliece, R. J.

- 42-105 Modifications of the Griesmer Bound, pp. 41-44.

G. Solomon

- 42-107 Some Partial-Unit-Memory Convolutional Codes, pp. 57-72.

See Abdel-Ghaffar, K.

Mellstrom, J. A.

- 42-105 Elevation Control System Model for the DSS 13 Antenna, pp. 83-108.

See Gawronski, W.

- 42-106 Modeling and Simulations of the DSS 13 Antenna Control System, pp. 205-248.

See Gawronski, W.

- 42-106 Pattern-Recognition Techniques Applied to Performance Monitoring of the DSS 13 34-Meter Antenna Control Assembly, pp. 30-51.

P. Smyth

- 42-108 Real-Time Antenna Fault Diagnosis Experiments at DSS 13, pp. 96-108.

C. Pierson and P. Smyth

Mizuhara, A.

- 42-108 250-kW CW Klystron Amplifier for Planetary Radar, pp. 201-213.

See Cormier, R.

Monk, A. M.

- 42-106 Carrier-to-Noise Power Estimation for the Block V Receiver, pp. 353-363.

Mueller, R. O.

- 42-105 Design and Analysis of a Low-Loss Linear Analog Phase Modulator for Deep Space Spacecraft X-Band Transponder Applications, pp. 136-145.

See Mysoor, N. R.

Mysoor, N. R.

- 42-105 Design and Analysis of a Low-Loss Linear Analog Phase Modulator for Deep Space Spacecraft X-Band Transponder Applications, pp. 136-145.

R. O. Mueller

Onyszchuk, I. M.

- 42-106 Coding Gains and Error Rates From the Big Viterbi Decoder, pp. 170-174.

- 42-106 Testing Interconnected VLSI Circuits in the Big Viterbi Decoder, pp. 175-182.

Ortiz, G. G.

- 42-106 A Cooled 1- to 2-GHz Balanced HEMT Amplifier, pp. 52-59.

S. Padin

- 42-106 Performance of a High T_c Superconducting Ultra-Low Loss Microwave Stripline Filter, pp. 60-65.

See Bautista, J. J.

Otoshi, T. Y.

- 42-105 A Portable X-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation—Part II: Tests on the Antenna, pp. 54-68.

S. R. Stewart and M. M. Franco

- 42-106 **A Portable Ka-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation—Part I: Design and Ground Tests**, pp. 249–265.
S. R. Stewart and M. M. Franco
- 42-106 **A Portable Ka-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation—Part II: Tests on the Antenna**, pp. 266–282.
S. R. Stewart and M. M. Franco
- 42-106 **Efficiency Calibration of the DSS 13 34-Meter Diameter Beam Waveguide Antenna at 8.45 and 32 GHz**, pp. 283–297.
See Slobin, S. D.
- 42-107 **A Portable Ku-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation**, pp. 73–80.
S. R. Stewart and M. M. Franco
- 42-107 **A Proposed Far-Field Method for Frequency-Stability Measurements on the DSS-13 Beam-Waveguide Antenna**, pp. 81–87.
- 42-108 **The Electrical Conductivities of the DSS-13 Beam-Waveguide Antenna Shroud Material and Other Antenna Reflector Surface Materials**, pp. 154–164.
M. M. Franco and H. F. Reilly, Jr.
- Padin, S.**
- 42-106 **A Cooled 1- to 2-GHz Balanced HEMT Amplifier**, pp. 52–59.
See Ortiz, G. G.
- Peng, T. K.**
- 42-106 **A Carrier-Arraying Demonstration at Goldstone for Receiving Pioneer 11 Signals**, pp. 307–334.
See Pham, T. T.
- Petty, S.**
- 42-108 **DSN G/T_{op} and Telecommunications System Performance**, pp. 271–278.
See Stelzried, C.
- Pham, T. T.**
- 42-106 **A Carrier-Arraying Demonstration at Goldstone for Receiving Pioneer 11 Signals**, pp. 307–334.
M. K. Simon, T. K. Peng, M. H. Brockman, S. S. Kent, and R. Weller
- Pierson, C.**
- 42-108 **Real-Time Antenna Fault Diagnosis Experiments at DSS 13**, pp. 96–108.
See Mellstrom, J.
- Pollara, F.**
- 42-105 **Integer Cosine Transform for Image Compression**, pp. 45–53.
See Cheung, K.-M.
- Posner, E. C.**
- 42-107 **Expected Antenna Utilization and Overload**, pp. 104–112.
- Prestage, J. D.**
- 42-108 **A High-Performance Hg⁺ Trapped Ion Frequency Standard**, pp. 10–18.
R. L. Tjoelker, G. J. Dick, and L. Maleki
- Rafferty, W.**
- 42-108 **DSN G/T_{op} and Telecommunications System Performance**, pp. 271–278.
See Stelzried, C.
- Randolph, L. W.**
- 42-105 **Orbital Debris Radar Instrumentation**, pp. 146–148.
See Goldstein, R. M.
- Rayhrer, B.**
- 42-106 **Design and Status of the RF-Digitizer Integrated Circuit**, pp. 155–159.
B. Lam, L. E. Young, J. M. Srinivasan, and J. B. Thomas

Reilly, H. R., Jr.

- 42-107 The L/C-Band Feed Design for the DSS 14 70-Meter Antenna (Phobos Mission), pp. 88-95.

See Stanton, P. H.

- 42-108 The Electrical Conductivities of the DSS-13 Beam-Waveguide Antenna Shroud Material and Other Antenna Reflector Surface Materials, pp. 154-164.

See Otoshi, T. Y.

Remer, D. S.

- 42-106 Modeling Preparation Costs for Space Missions by Using Major Cost Drivers, pp. 404-415.

See Sherif, J. S.

Rinker, G. C.

- 42-108 The X-Windows Interactive Navigation Data Editor, pp. 66-78.

Rochblatt, D. J.

- 42-108 A Microwave Holography Methodology for Diagnostics and Performance Improvement for Large Reflector Antennas, pp. 235-252.

- 42-108 Performance Improvement of DSS-13 34-Meter Beam-Waveguide Antenna Using the JPL Microwave Holography Methodology, pp. 253-270.

B. L. Seidel

Rogstad, D. H.

- 42-107 Suppressed Carrier Full-Spectrum Combining, pp. 12-20.

Sadr, R.

- 42-106 The Advanced Receiver II Telemetry Test Results at Goldstone, pp. 119-131.

R. Bevan and S. Hinedi

Schwartz, J. A.

- 42-106 Design and Testing of an Active Quenching Circuit for an Avalanche Photodiode Photon Detector, pp. 102-110.

See Arbel, D.

Seidel, B. L.

- 42-108 Performance Improvement of DSS-13 34-Meter Beam-Waveguide Antenna Using the JPL Microwave Holography Methodology, pp. 253-270.

See Rochblatt, D. J.

Shah, B.

- 42-106 Symbol Lock Detection in the ARX II and Block V Receivers, pp. 132-154.

See Shihabi, M.

Shahshahani, M.

- 42-105 Integer Cosine Transform for Image Compression, pp. 45-53.

See Cheung, K.-M.

- 42-107 A Comparison of the Fractal and JPEG Algorithms, pp. 21-26.

See Cheung, K.-M.

Shaik, K.

- 42-106 Spectral Filters for Laser Communications, pp. 93-101.

Shell, J.

- 42-106 Equivalent Circuit Model of Traveling-Wave Maser Slow-Wave Structures, pp. 66-92.

Sherif, J. S.

- 42-106 Modeling Preparation Costs for Space Missions by Using Major Cost Drivers, pp. 404-415.

D. S. Remer and H. R. Buchanan

42-108 Software-Engineering Process Simulation Model (SEPS), pp. 165-185.

See Lin, C. Y.

Shihabi, M.

42-106 Symbol Lock Detection in the ARX II and Block V Receivers, pp. 132-154.

S. Hinedi and B. Shah

Simon, M. K.

42-106 A Carrier-Arraying Demonstration at Goldstone for Receiving Pioneer 11 Signals, pp. 307-334.

See Pham, T. T.

Slobin, S. D.

42-106 Efficiency Calibration of the DSS 13 34-Meter Diameter Beam Waveguide Antenna at 8.45 and 32 GHz, pp. 283-297.

T. Y. Otoshi, M. J. Britcliffe, L. S. Alvarez, S. R. Stewart, and M. M. Franco

Smyth, P.

42-106 Pattern-Recognition Techniques Applied to Performance Monitoring of the DSS 13 34-Meter Antenna Control Assembly, pp. 30-51.

See Mellstrom, J.

42-108 Real-Time Antenna Fault Diagnosis Experiments at DSS 13, pp. 96-108.

See Mellstrom, J.

Solomon, G.

42-105 Modifications of the Griesmer Bound, pp. 41-44.

See McEliece, R. J.

42-107 Some Partial-Unit-Memory Convolutional Codes, pp. 57-72.

See Abdel-Ghaffar, K.

42-108 Nonlinear, Nonbinary Cyclic Group Codes, pp. 84-95

Sovers, O. J.

42-106 JPL 1990-3: A 5-nrad Extragalactic Source Catalog Based on Combined Radio Interferometric Observations, pp. 364-383.

Sridhar, S.

42-106 Performance of a High T_c Superconducting Ultra-Low Loss Microwave Stripline Filter, pp. 60-65.

See Bautista, J. J.

Srinivasan, J. M.

42-106 Design and Status of the RF-Digitizer Integrated Circuit, pp. 155-159.

See Rayhrer, B.

Stanton, P. H.

42-106 Wideband Waveguide Polarizer Development for SETI, pp. 384-393.

See Lee, P.

42-107 The L/C-Band Feed Design for the DSS 14 70-Meter Antenna (Phobos Mission), pp. 88-95.

H. R. Reilly, Jr.

Stelzried, C.

42-108 DSN G/T_{op} and Telecommunications System Performance, pp. 271-278.

R. Clauss, W. Rafferty, and S. Petty

Stewart, S. R.

42-105 A Portable X-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation—Part II: Tests on the Antenna, pp. 54-68.

See Otoshi, T. Y.

42-106 A Portable Ka-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation—Part I: Design and Ground Tests, pp. 249-265.

See Otoshi, T. Y.

- 42-106 A Portable Ka-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation—Part II: Tests on the Antenna, pp. 266–282.
See Otoshi, T. Y.
- 42-106 Efficiency Calibration of the DSS 13 34-Meter Diameter Beam Waveguide Antenna at 8.45 and 32 GHz, pp. 283–297.
See Slobin, S. D.
- 42-107 A Portable Ku-Band Front-End Test Package for Beam-Waveguide Antenna Performance Evaluation, pp. 73–80.
See Otoshi, T. Y.
- 42-108 DSS-13 Beam-Waveguide Antenna Performance in the Bypass Mode, pp. 135–153.
- 42-108 Performance Improvement of DSS-13 34-Meter Beam-Waveguide Antenna Using the JPL Microwave Holography Methodology (Appendix)
See Rochblatt, D. J.
- Swanson, L.
- 42-106 Frame Error Rate of the NASA Concatenated Coding System, pp. 183–187.
K.-M. Cheung
- 42-108 Node Synchronization Schemes for the Big Viterbi Decoder, pp. 186–200.
See Cheung, K.-M.
- Thomas, J. B.
- 42-106 Design and Status of the RF-Digitizer Integrated Circuit, pp. 155–159.
See Rayhrer, B.
- Thurman, S. W.
- 42-107 Trajectory and Navigation System Design for Robotic and Piloted Missions to Mars, pp. 113–131.
S. E. Matousek
- 42-108 Orbit-Determination Performance of Doppler Data for Interplanetary Cruise Trajectories, Part I: Error Analysis Methodology, pp. 31–48.
See Ulvestad, J. S.
- Tjoelker, R. L.
- 42-108 A High-Performance Hg⁺ Trapped Ion Frequency Standard, pp. 10–18.
See Prestage, J. D.
- Ulvestad, J. S.
- 42-105 Phasing the Very Large Array on Galileo in the Presence of Jupiter's Strong Radio Emission, pp. 69–82.
- 42-108 Orbit-Determination Performance of Doppler Data for Interplanetary Cruise Trajectories, Part I: Error Analysis Methodology, pp. 31–48.
S. W. Thurman
- 42-108 Orbit-Determination Performance of Doppler Data for Interplanetary Cruise Trajectories, Part II: 8.4-GHz Performance and Data-Weighting Strategies, pp. 31–48.
- Veruttipong, W.
- 42-105 Gaussian Beam and Physical Optics Iteration Technique for Wideband Beam Waveguide Feed Design, pp. 128–135.
J. C. Chen and D. A. Bathker
- Vijayaraghavan, A.
- 42-108 Position Determination of a Lander and Rover at Mars With Earth-Based Differential Tracking, pp. 279–293.
See Kahn, R. D.
- Vilnrotter, V. A.
- 42-106 Channel Assignments and Array Gain Bounds for the Ka-Band Array Feed Compensation System, pp. 160–169.

Wang, R. T.

- 42-107 **A Receiver Design for the Superconducting Cavity-Maser Oscillator**, pp. 1-11.

G. J. Dick

Weller, R.

- 42-106 **A Carrier-Arraying Demonstration at Goldstone for Receiving Pioneer 11 Signals**, pp. 307-334.

See Pham, T. T.

Yan, T.-Y.

- 42-106 **Analysis of a Multiple Reception Model for Processing Images From the Solid-State Imaging Camera**, pp. 111-118.

Young, L. E.

- 42-106 **Design and Status of the RF-Digitizer Integrated Circuit**, pp. 155-159.

See Rayhrer, B.

Zahopoulos, C.

- 42-106 **Performance of a High T_c Superconducting Ultra-Low Loss Microwave Stripline Filter**, pp. 60-65.

See Bautista, J. J.

Zimmerman, G. A.

- 42-107 **A 640-MHz 32-Megachannel Real-Time Polyphase-FFT Spectrum Analyzer**, pp. 132-140.

M. F. Garyantes, M. J. Grimm, and B. Charny

- 42-107 **Polyphase-Discrete Fourier Transform Spectrum Analysis for the Search for Extraterrestrial Intelligence Sky Survey**, pp. 141-154.

S. Gulkis